Application of time-series analysis for a cable-stayed bridge monitoring using GPS long-term data

By

Hien Van Le

Dissertation

Submitted in partial fulfillment of requirements
for the degree of Doctor of Philosophy in Engineering
in the Graduate School of Urban Innovation of
YOKOHAMA National University, 2015

Doctoral Committee

Associate Professor Mayuko NISHIO
Professor Hitoshi YAMADA
Professor Hiroshi KATSUCHI
Professor Yozo FUJINO
Professor Kazuo KONAGAI
Doctor Dionysius SIRINGORINGO

Abstract

The Global Positioning System (GPS) technology has been successfully used to measure displacements of oscillating flexible civil engineering structures, such as suspension bridges and high-rise buildings. One of the important advantages of the GPS technology is that it can reach to the millimeter accuracy level in displacement measurements. The global deformation of structures can be understood based on GPS data. However, similar to other methods in Structural Health Monitoring (SHM) system, the GPS technology provides a huge amount of data in continuous monitoring of structures. Thus, how to convert a huge amount of data into usable information, and how to acquire the long-term static displacements of large-scale structures with stable and high accuracies are important tasks to study in GPS data processing. There was no strategy for GPS long-term data analysis in SHM field. The main purposes of this study are to deal with these shortcomings.

Furthermore, some previous researchers studied on the SHM system of structures which pointed out that the long-term monitoring data were greatly affected by the environmental and operational effects. In this case of study, the correlation analysis is applied to analyze the effects of environmental factors, such as temperature, to monitoring data. The time series analysis is also useful method that would be applied to extract the usable information from the monitoring data. The purpose of this thesis focuses on two main parts. First of all is to process the GPS time series data of a long-span bridge. The Autoregressive Integrated Moving Average (ARIMA) model is applied to extract the usable information of the target bridge that can be used to assess the structural conditions. The AR-MA coefficients which extracted from the ARIMA model estimation are used as the multivariate features for comparing by applying the Mahalanobis distance method. In the second purpose of this thesis, a numerical analysis of a cable-stayed bridge model under thermal change is conducted to test the effects of thermal change to the global deformation of the bridge model. The ARIMA model is also applied for the simulated displacement data to extract the coefficients. The effects of thermal change to global deformation of bridge model which can be realized by the distributions of the AR-MA coefficients. The results of present study propose that the extracted AR-MA coefficients are effective features to show the structural conditions, and application of time-series analysis is useful strategy to process the long-term monitoring data of structures.

The present dissertation is consisted of six chapters, and the contents of all chapters are summarized as follows:

Chapter 1 introduces the status of long-span bridge constructions around the world. The role of the structural health monitoring system to assess the structural conditions is also discussed. The general GPS technology is introduced in the later part of this chapter as well as the capability of the GPS application in the SHM system is mentioned. Then, the necessity of this research is presented. The aims, objectives and methodology are also defined in this chapter.

Chapter 2 explains the target bridge of present study as well as the GPS system used in the target bridge. The observed GPS data are analyzed to deal with missing data, and the correlation between GPS time-history data and temperature data is also studied to show the global deformation of the target bridge.

Chapter 3 explains the theory of the time-series Autoregressive Integrated Moving Average (ARIMA) model and its application to process the GPS data.

Chapter 4 presents the numerical of a cable-stayed bridge model due to thermal changes. The simulated displacement data acquired from the finite element analysis due to thermal changes are applied the ARIMA model to extract the AR-MA coefficients for assess the structural conditions.

Chapter 5 presents the statistical condition assessment of the practical target bridge that using the AR and MA coefficients gained from chapter 3 as well as the numerical coefficients derived from chapter 4. In order to study in this chapter, the AR and MA coefficients are used as the multivariate features that are compared by using the Mahalanobis distance method.

Chapter 6 gives some conclusions based on the results of current study. The perspective of the future works is also recommended in this chapter.

Acknowledgements

First and foremost, I would like to thank from my deepest and most sincere gratitude to my advisor, Associate Professor Mayuko Nishio, for her supports and encouragement throughout my study. She spent a lot of time to teach me, and her mentorship has been critical for the success of my work. I will forever grateful to her, and I hope to cooperate more with her in future researches.

I would also like to say many thanks and all gratitude to Professor Hitoshi Yamada, for his great helps. He also taught me so much during his courses as well as giving me the useful advices to accomplish this dissertation. Sincerely, I am always grateful to him for his effort to open opportunities for young scientists coming from Vietnam.

Next, I would sincerely like to thank to Professor Hiroshi Katsuchi for his advices and recommendations in my study. In my deepest, my work would not to be accomplished without his great helps.

Additionally, I would also like to thank to all structure laboratory members. I had wonderful moments to be studied with you together. I really enjoyed the life in Japan through your guidance.

Sincerely, I would like to say thank from my deepest to Dr. Ho Lan Huong and all my colleagues, friends in Vietnam for their ongoing support and encouragement in my work.

Last, but not least, with all my sincere gratitude which I would like to thank to my families, my dear wife and lovely son who always support me. To my dear wife, thank you so much for your love, encouragement and your take care of our lovely son.

Hien Van Le

List of contents

List of figures

List of tables

Nomenclature

Chapter 2

a_i	Polynomial parameters
t_i	Ordinal number of point i
x_i	Known data
m	Polynomial order
n	Number of data points
$\mathbf{M}$	Matrix of ordinal numbers
$\mathbf{x}$	Known data vector
$\mathbf{a}$	Polynomial parameter vector
r	Correlation coefficient
μ	Mean
σ	Standard deviation
E	Expectation

Chapter 3

AR	Autoregressive
MA	Moving Average
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
X_t	A series
ϕ_i	Autoregressive coefficients
θ_i	Moving Average coefficients
ε_i	White noise
p	Order of autoregressive term
q	Order of moving average term
d	Order of integrated term
B	Backward shift operator
c	Constant
∇	Differencing operator
ρ_i	Autocorrelation patterns
BIC	Bayesian Information Criterion
DF_τ	Dickey-Fuller test

$\hat{\gamma}$	Least-square estimate
$SE(\hat{\gamma})$	Usual standard error
Q	Ljung-Box Q test

Chapter 4

$temp$	Temperature data
t	Time interval
A	Temperature amplitude
SD_z	Stiffness direction along to z direction
SR_y	Stiffness rotation along to y direction
BC	Boundary condition change
TC	Tension change

Chapter 5

MD	Mahalanobis distance
$\mathbf{x}_i$	Multi-dimensional feature vector
$\bar{\mathbf{x}}$	Mean vector
$\mathbf{S}$	Covariance matrix
MD^B	Mahalanobis distance of base
$\mathbf{x}^B$	Multi-dimensional vector of base
$\mathbf{S}_B$	Covariance matrix of base
$\bar{\mathbf{x}}^B$	Mean vector of base
$\mathbf{x}^E$	Multi-dimensional vector of estimation
MD^E	Mahalanobis distance of estimation

Chapter 1

Introduction

1.1. Background

1.1.1. Bridge constructions and Structural Health Monitoring (SHM) of bridges

In most countries, the weight of new civil engineering constructions is constantly decreasing to the benefit of structure maintenance, basically for economic reasons. In the meantime, stakeholders are concerned about obtaining a precise knowledge of the health of the existing infrastructures, their preservation and the evolution of their safe-keeping. Bridge is one of the important infrastructures in the national economy, which are considered as the crucial link in transport network. Recently, many kinds of long-span bridges have been building in many countries around the world which are called the suspension bridge and the cable-stayed bridge. One of the most famous long-span bridges can be listed as Akashi Kaikyo bridge in Japan which is known as the longest span (1991meter of length) at time of completion in 1998(Figure 1.1). The Tsing Ma bridge was constructed in Hong Kong in 1997 which is 1337 meter of main span length as the world's ninth-longest span suspension bridge, and was the second longest at time of completion (Figure 1.2).

Figure 1.1.Akashi Kaikyo bridge [1]

Figure 1.2.Tsing Ma bridge [2]

However, there were many bridge failures which were caused by normal or abnormal loadings during their operations. One of the famous disasters in the past of civil engineering field was known as the collapse of Tacoma Narrows bridge in America in 1940 which sparked renewed

research into the aerodynamics of the suspension bridge (Figure 1.3). Recently, the collapse of the I-35W Bridge in Minearsota, America, in 2007, which have raised many concerns regarding to the current condition of bridges (Figure 1.4).

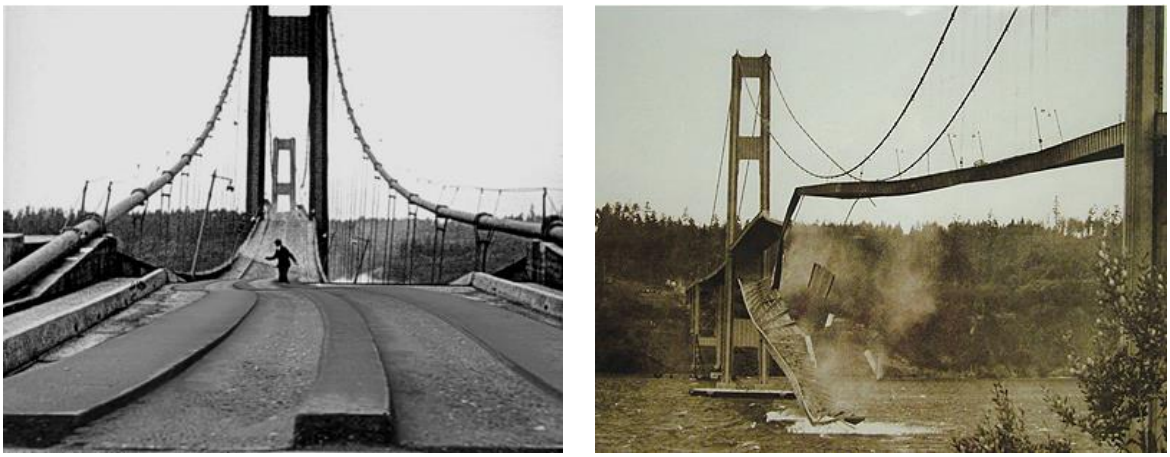


Figure 1.3. The collapse of the Tacoma Narrows bridge [3]

Figure 1.4. The I-35W Bridge collapse [4]

Basically, the Structural Health Monitoring (SHM) aims to give a diagnosis of the "state" of the constituent materials, of the different parts, and of the full assembly of these parts constituting the structures as a whole, at every moment during the life of a structure. The principle and organization of SHM system can be summarized Figure 1.5 [5]. The SHM plays an important role which has been using in many fields of structural monitoring such as aerospace engineering, civil engineering etc. In structural health dynamic monitoring field, SHM can be applied to provide valuable information in current evaluation of structure integrity, durability and reliability.

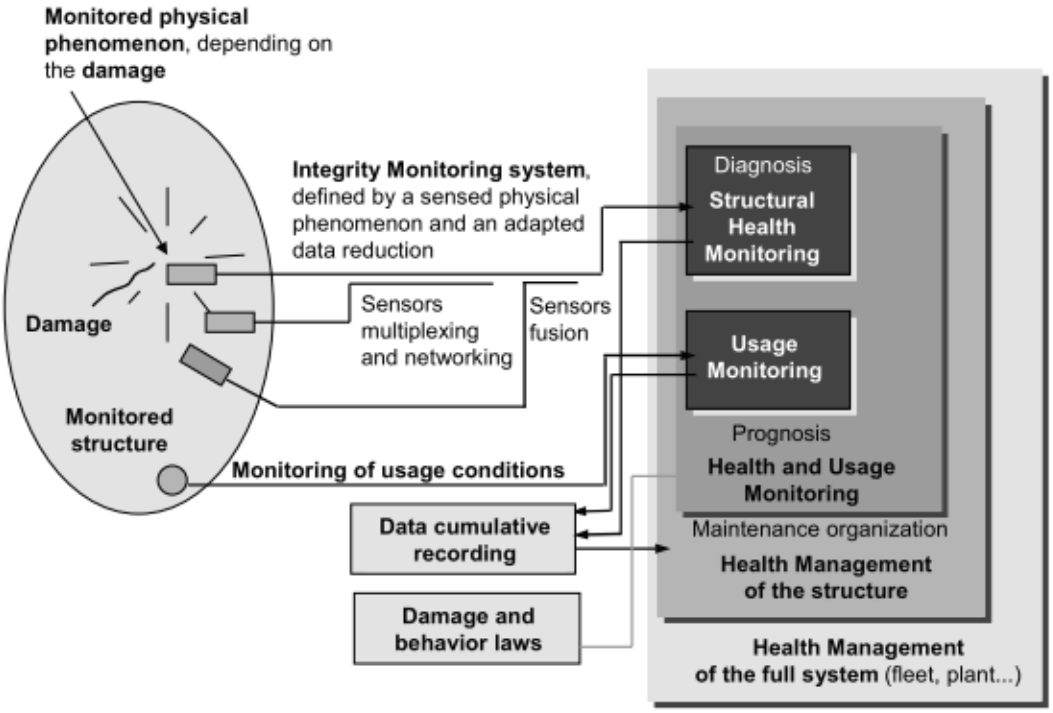


Figure 1.5. Principle and organization of a SHM system [5]

Furthermore, the structural health monitoring (SHM) has been considering for the process of implementing a structural condition assessment for the large-scaled structures, such as long-span bridges. In those flexible bridges, some structural changes have been considered to relate closely to deformations [6]. A bridge usually has two kinds of individual deformations: long-term and short-term deformations. The long-term deformations are irrecoverable or periodic; they are caused by the foundation settlement, the creep, the temperature effects, and so on. The short-term deformations are caused by dynamic inputs, such as those induced by wind, tidal current, earthquake, and traffic [6-8]. The monitoring of those deformations is thus expected to be appropriate to capture the structural changes.

In research field of SHM, some papers studied on the continuous structural monitoring data, in most of which, the dynamic characteristics were used to verify the structural changes [9-11]. Those studies pointed out that the long-term monitoring data were greatly affected by the environmental and operational effects. Sohn *et al* [9] mentioned that those effects consisted of temperature, humidity, and the changes in operational loads and boundary conditions. The variability of monitoring data due to the environmental effects could then mask more subtle structural changes caused by damages. In their study, a linear adaptive filter model was examined

to discriminate the changes of modal parameters due to temperature changes from those caused by structural damage or other environmental effects. The results indicated that a linear adaptive filter to temperature could reproduce the natural variability of the frequencies with respect to time of a day. Cornwell *et al* [10] studied on the variability in modal parameters due to the environmental effects and the operational conditions. In their study, the correlation analysis was conducted using the resonant frequencies and the temperature data measured in twice, and the high correlation coefficients among them indicated the high influences of the temperature changes on the changes of structural properties. Farrar *et al* [11] studied on quantifying the variability in identified modal parameters caused by some sources, such as variability in testing procedures, in test conditions, and the environmental variability. Most of those studies then concluded that the consideration of the variability in monitoring data due to the environmental and operational effects was a requirement for effective structural health monitoring.

However, the SHM system still remains some issues required to be considered, such as how to acquire the long-term static displacements of large-scale structures with stable and high accuracies, and how to convert large amount of data into usable information. In most cases of using SHM, the major devices to capture dynamic/static structural responses are the accelerometer and the strain sensors. The acceleration measurement is the standard method to measure the structure vibration. Accelerometer sensor is very small size and light, therefore, the effects of this kind of sensor to the properties of the vibration system is minimal. However, we have to use wiring to connect directly between sensors and structures as well as between sensors and the central recording unit. Moreover, in most case of papers as mentioned, the long-term deformation of bridge monitoring which were only considered in the local site. Therefore, the global deformation could not realize.

On the other hand, there are some sensors for measuring the structural displacements, such as the laser interferometer and some electronic distance instruments. Although those sensors have the advantage of high accuracies; however, they also have some disadvantages; e.g., it is difficult to capture the measurement points when the displacement becomes too large, it is difficult to acquire data in real time, and the data acquisition is limited by climate conditions; i.e., the clear line of sight is one of basic requirements [12].These disadvantages can be solved by the Global Positioning System method which satisfies the accurate requirements of SHM, and has several advantages in its applications.

1.1.2. Global Positioning System introduction

The Global Positioning System (GPS) is a satellite-based navigation system that was developed by the U.S. Department of Defense (DoD) in the early 1970s. Initially, GPS consists of a constellation of 24 operational satellites which are known as the initial operational capability (IOC), was completed in July 1993 [14]. The satellite altitude is about 20,200 km above the Earth's surface. To ensure continuous worldwide coverage, GPS satellites are arranged so that four satellites are placed in each of six orbital planes (Figure 1.6). With this constellation geometry, four to ten GPS satellites will be visible anywhere in the world, if an elevation angle of 10^0 is considered. To provide the positioning, or location, information, only four satellites are needed. Nowadays, the number of satellites on the space which is more than 30, therefore, there are always more than four visible satellites anywhere in the world that can give more accurate for measurements.

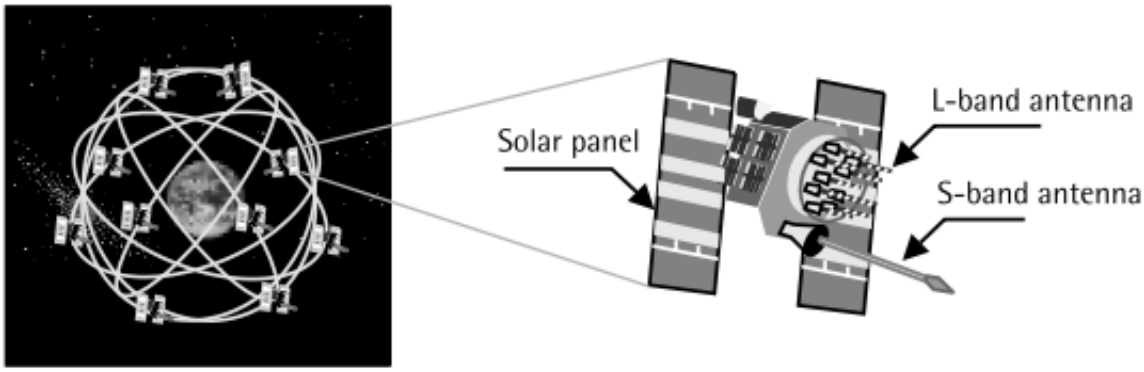


Figure 1.6. GPS constellation [14]

GPS consists of three segments: the space segment, the control segment, and the user segment (Figure 1.7). The space consists of the satellite constellation which each GPS satellite transmit a signal. The control segment of the GPS system consists of a worldwide network of tracking stations. The primary task of the operational control segment is tracking the GPS satellites in order to determine and predict satellite locations, system integrity, behavior of the satellite atomic clocks etc. The user segment includes all military and civilian users. With a GPS receiver connected to a GPS antenna, a user can receive the GPS signals, which can be used to determine its position anywhere in the world. The idea behind GPS is rather simple. If the distances from a point on the Earth (a GPS receiver) to three GPS satellites are known along with

the satellite locations, then the location of the point (or receiver) can be determined by simply applying the well-known concept of resection (Figure 1.8).

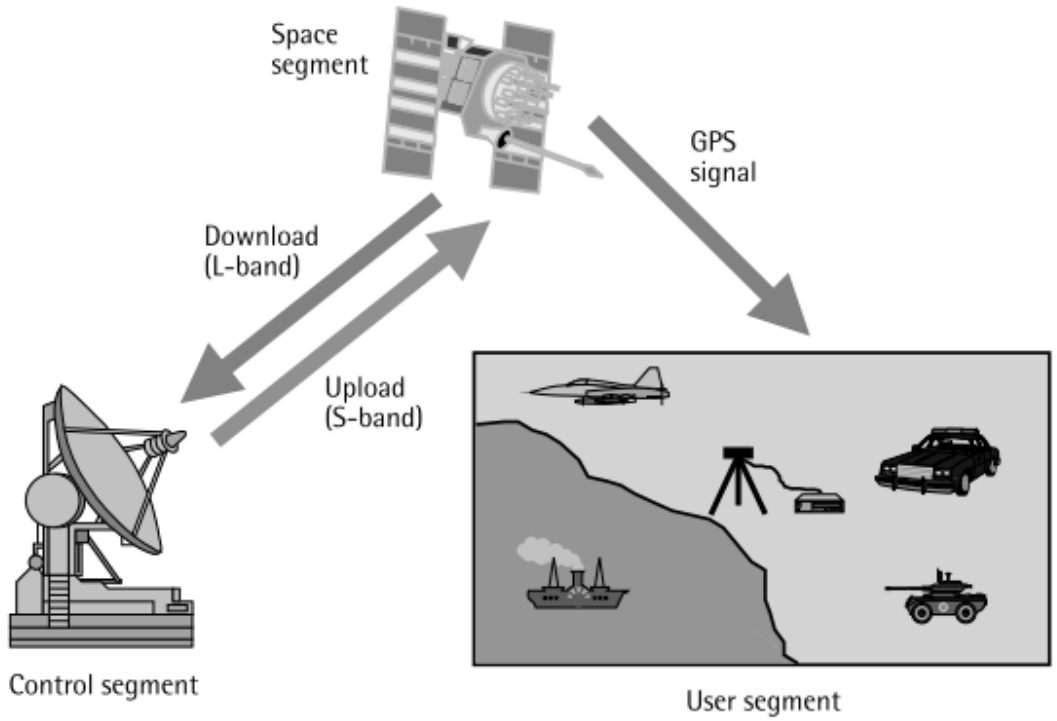


Figure 1.7. GPS segments [14]

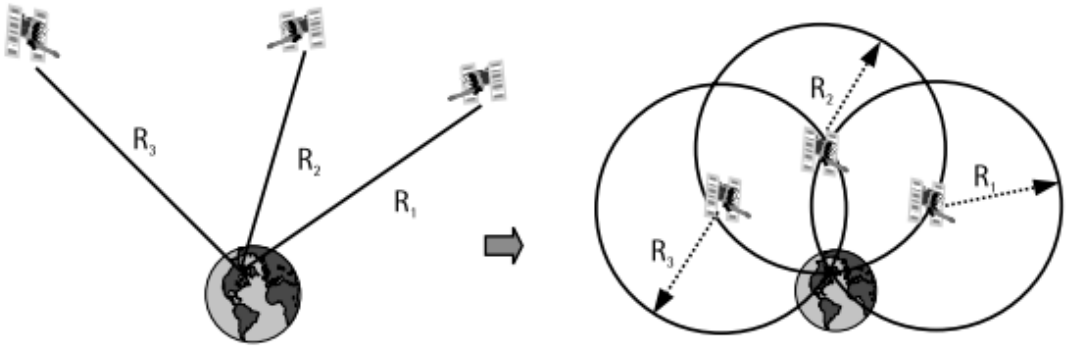


Figure 1.8. Basic idea of GPS positioning [14]

For using GPS technology in the SHM system which requires the high-accuracy, the Real-Time Kinematic GPS (RTK-GPS) technique is used. The origin of RTK-GPS technique is based on the Differential GPS (DGPS) technique that data from a receiver at a known location is used to correct the data from other one at an unknown location. Differential corrections can be

applied in either real-time or by post-processing. Since most of the errors in GPS are common to users in a wide area, the DGPS-corrected solution is significantly more accurate than a normal GPS solution [12].

1.1.3. GPS monitor system composing in SHM

A GPS monitor system of structural monitoring is consisted of GPS reference station (base station), remote station (rover station), monitoring center and communication system as summarized in Figure 1.9, and how a GPS monitor system works on the case of bridge monitoring is summarized in Figure 1.10. In the monitor system, a reference station has a task to track all GPS satellites that it can observe in its location, and all antennas in remote stations simultaneously received satellite signals from several satellites. To define coordinates of one point in three-dimensional directions, at least four satellites are required to obtain accurate positioning. At the reference station, the different information received from the satellites is transmitted to the remote stations in real time by using the cable of the communication system. In the remote station, both signals received from the satellites and the different information from the reference station is received, and the remote station coordinates in three-dimensional can be calculated by carry out the real-time different processing using GPS software. The coordinates are then transmitted to the GPS monitoring center, and are further processed to get the structural displacements, rotation angle in a special direction and so on. All displacements are stored into database to make health assessment about the structures in future. Notice that, the distance between the reference station and the remote station relates to the effects of the atmospheric and orbital errors whereas the short distance between them is appreciated to set up. At present case, when the distance was one kilometer or less, the effects of atmospheric and orbital errors are expected to be at or below the millimeter level [12].

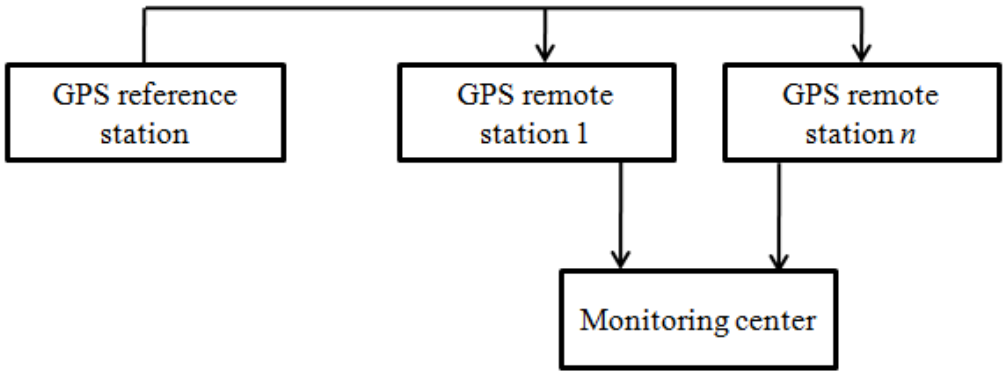


Figure 1.9. Graph of GPS monitor system composing [12]

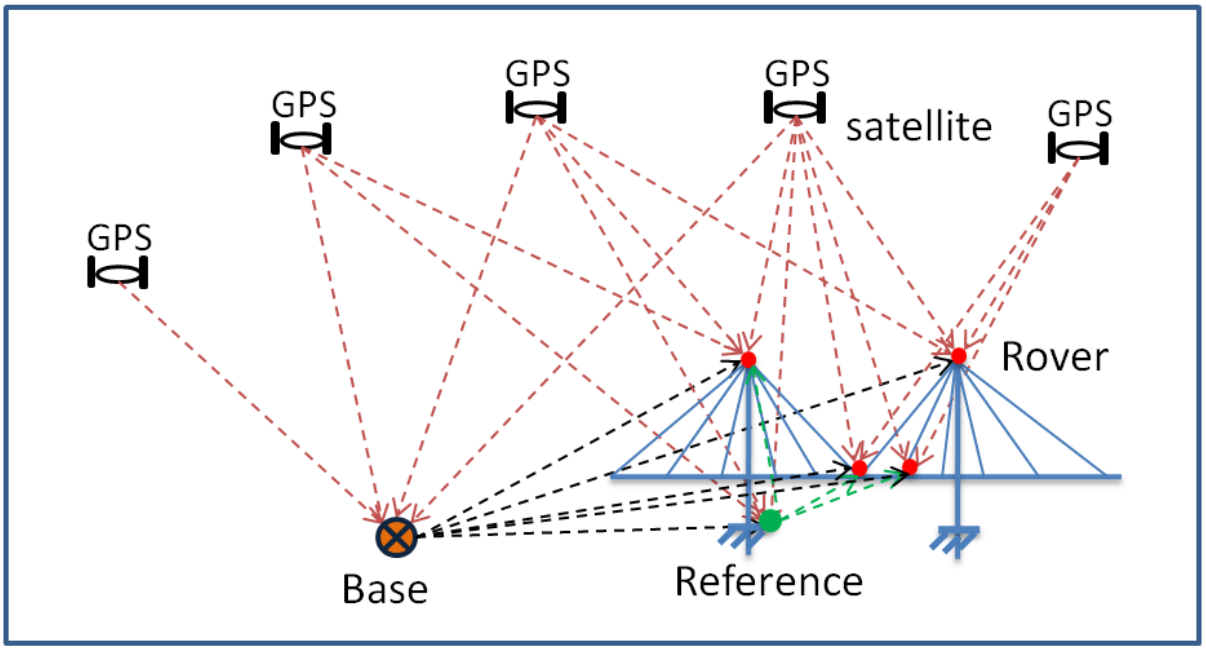


Figure 1.10. RTK-GPS on bridge monitoring

In addition, the selection of the remote station is related to the monitoring task, deformation size, and the captured information of the structures. Generally, the remote station should be located at the interest points, the biggest deformation points or the feature points of structures. For instance, in the long-span bridges monitoring, the remote station are located at the middle of bridge deck, on fourth of the bridge deck or on the top of towers.

The Global Positioning System (GPS) technology has been successfully used to measure displacements of oscillating flexible civil engineering structures, such as suspension bridges and high-rise buildings. There are some advantages of the GPS technology to monitor the displacements of the large-scaled civil structures; e.g., it overcomes the limitation of climate, it can also measure the structural displacements in the three-dimensional directions at millimeter level accuracies, and most GPS monitoring system operates automatically [6,12]. Moreover, the accuracy of GPS technology is improving, the sensor is reducing in weight and size as well as the system is more users friendly. Therefore, the GPS technology can be applied for continuous monitoring in SHM system of the long-span bridges. The study of the GPS applications in SHM monitoring of long-span bridges is necessary to assess the structural conditions as well as health structures during their operation.

Actually, there are some cases of the SHM using GPS technology for the monitoring of long-span bridges; e.g., the *Tianjin Yonghe* cable stayed bridge in Hong Kong [6], the *Akashi*

Kaikyo Bridge in Japan [13]. In the case of the *Akashi Kaikyo* bridge which is a suspension bridge with the center span of 1,991m, the GPS system has been installed to measure the displacements in three directions at three locations; the center span, the top of one of towers and the anchorage. In the indicated study in [13], the configuration of a whole suspension bridge was investigated based on the statistical methods that were applied for the acquired long-term GPS data. They showed that the location when deformation is caused by an earthquake or typhoon which can be identified by using GPS data. Mosbeh R. Kaloop *et al* (2009) studied the Tianjin Yonghe cable-stayed bridge monitoring using GPS technique for deformation data collection [7]. In their study, two analytical methods were used namely Kalman Filter and Parametric Least Square for the adjustment of the GPS data. Their results showed that by using Real Time Kinematic technique in monitoring bridge, the deformation of the interested points could be carry out; they found that the traffic loads are the main factor affecting bridge cracks. However, there were only two GPS remote sensors were located on the top of towers in their target bridge. Thus, it could not to assess the global deformation of bridge by GPS data. The effects of environment to the target bridge deformations did not mentioned in this study.

In case of GPS data processing in SHM, a huge amount of static time-series data is acquired by monitoring system. Similar to other methods, a question is that how to convert large amount of GPS time-series data into usable information which needs to solve. There were then actually some studies that used the time-series analysis to assess the structural conditions from the monitoring data that consisted of the environmental and operational effects. Sohn *et al* [15] statistically examined the changes in the autoregressive (AR) model coefficients that estimated from dynamic data. It was shown that the distributions of AR coefficients estimated from data sets, which were from undamaged and damaged systems, could be appropriately classified to exact conditions. Ruigen Yao *et al* [48] used AR model for statistical pattern recognition algorithms in both simulation and experimental test which showed the ability of using AR coefficient for damage detection of structures. Additionally, Omenzetter *et al* [16] used a seasonal autoregressive integrated moving average model with exogenous inputs (SARIMAX) and transfer function to model the relationship between strain data and temperature data. In this study, unusual structural condition changes or damages could be detected by applying the outlier detection and the intervention analysis technique to the estimated model. In the other study from the same authors [17], the approach to apply a vector seasonal autoregressive integrated moving average (ARIMA) time-series model was also presented. This study showed that the coefficients of the

ARIMA model that were estimated by the adaptive Kalman filter could also be used for detecting the unusual events occurred on the structure.

In most case of studies in SHM system as mentioned, the time series analysis was applied for the static local strain data which could only assess the structural conditions in the local deformation site. In this current study, the new points are that the static GPS time series data are analyzed to get the global deformation of a long span bridge, and the extracted features are used for the statistical comparison by applying the Mahalanobis distance. Furthermore, the sensitivity of global deformation of model due to thermal changes is also studied in numerical analysis.

1.2. Objectives and methodology of research

The aims of current study is to get better understanding of the application of GPS in the SHM system of the long-span bridge as well as to analyze the long-term displacement data acquired from a GPS monitoring system in a long-span cable-stayed bridge. The target bridge here is the *Can Tho* bridge, which is one of cable-stayed bridges in Vietnam. The time-series analysis method was applied to the GPS observed data collecting from the target bridge. We investigated the global deformation patterns mainly due to the temperature effects, and verified whether they can be used as the structural response features for the statistical structural condition assessment. On the other hand, the current study is also to present the sensitivity study using the simulated displacement data of a cable-stayed bridge model acquired from the finite element analysis to investigate the effects of temperature change to the global deformation of bridge model by applying the time-series Autoregressive Integrated Moving Average (ARIMA) model. In order to process the GPS data of the target bridge, the MATLAB software was mainly used.

According to the range of aims mentioned above, the objectives of this current dissertation are listed as follows:

- The quality of obtained GPS data is discussed and the missing data are handled by applying the least-squares approximation to interpolate missing values.

- The correlation coefficient analysis is conducted using the GPS data and the air temperature data to investigate the global deformation modes due to the temperature effects.

- To verify the applicability of one of the time-series models called ARIMA model for using the global deformations for the statistical structural condition assessment.

- The estimated AR and MA coefficients are used as the features of the target bridge for comparing by using Mahalanobis distance method. The results of comparison can be used for the structural condition assessment.

- A normal cable-stayed bridge model is built in the finite element analysis to investigate the effects of thermal changes to the global deformation of model. An ARIMA model is then applied for numerical data to assess the structural conditions.

1.3. Outline of the dissertation

The present dissertation is consisted of six chapters, and the list of all reference is added in the end of dissertation. The contents of all chapters are summarized as follows:

Chapter 1 introduces the status of long-span bridge constructions around the world. The role of the structural health monitoring system to assess the structural conditions is also discussed. The general GPS technology is introduced in the later part of this chapter as well as the capability of the GPS application in the SHM system is mentioned. Then, the necessity of this research is presented. The aims, objectives and methodology are also defined in this chapter.

Chapter 2 explains the target bridge of present study as well as the GPS system used in the target bridge. The observed GPS data are analyzed to deal with missing data, and the correlation between GPS time-history data and temperature data is also studied to show the global deformation of the target bridge.

Chapter 3 explains the theory of the time-series Autoregressive Integrated Moving Average (ARIMA) model and its application to process the GPS data.

Chapter 4 presents the numerical of a cable-stayed bridge model due to thermal changes. The simulated displacement data acquired from the finite element analysis due to thermal changes are applied the ARIMA model to assess the structural conditions.

Chapter 5presents the statistical condition assessment of the target bridge that using the AR and MA coefficients gained from chapter 3as well as the numerical coefficients derived from

chapter 4. In order to study in this chapter, the AR and MA coefficients are used as the multivariate features that are compared by using the Mahalanobis distance method.

Chapter 6 gives some conclusions based on the results of current study. The perspective of the future works is also recommended in this chapter.

Chapter 2

Target bridge introduction and GPS observed data analysis

2.1. Target bridge and installed GPS system

The target structure in this study was the *Can Tho* bridge, which is the longest cable-stayed bridge in the South East Asia opened in 2010. Figure 2.1 shows the location and a picture of the target bridge. It is the bridge over the *Hau* river in the south of Vietnam, with the total length of 2,750m, the center span of 550m, and the height of towers are 171m. The bridge has basically a concrete box-girder with the width of 26m; however, to increase the loading capacity, apart of the center span (middle 210m length) is made by a steel box-girder. The SHM system has been installed since 2010, which includes not only the GPS system but also many sensors, such as temperature sensors, anemometers, and accelerometers etc.

(a). Location of bridge [18] (b). The bridge photo [19]

Figure 2.1 The target bridge

The Figure 2.2 indicates the monitored sensor arrangement in the design stage of the target bridge where it plans to apply many kinds of monitored sensors such as anemometer, thermometer and so on. The planning of sensor system arrangement in the design stage was also shown in detail as indicated in Table 2.1 [20]. The GPS system installed in the bridge consists of nine sensors as the rover stations and two base stations as shown in Figure 2.3. The rover stations were

placed on the top of two towers, the center span of the girder (the upstream and downstream sides), the quarter of the center span (the upstream and downstream sides), and on the top of piers. One of two base stations was placed on the pile cap of the tower crane located on the southeast side of the North tower, while the other one was placed near the monitoring management office that was located one-kilometer far away from the southern side of the bridge. The adopted GPS equipments were the products of Leica co., ltd, GMX 902 GG model. The accuracies of the GPS system based on real-time kinematic technique are ±10mm ± 1ppm for the horizontal plane and ±20 mm ± 1ppm for the vertical direction. The data acquisition system was constructed, in which the GPS signal at each rover station was acquired in 20Hz, and their ten-minutes-averaged values were calculated. The averaged three-dimensional coordinates from the base station on the footing of the North tower were then save in each ten minutes; therefore, the data became time-series data with ten-minutes interval.

Table 2.1 Planning of the sensor system, network and wiring [20]

	Module of Hardware/Devices	Functions
1	Monitoring Office	. Monitoring and Recording various data . GPS station unit (1 unit) . Internet delivery
2	Top of the South Pylon	. GPS measurement unit (1 unit) . Anemometer (1 unit) . Monitoring camera (2 units) . Thermo-couples (4 units in the pylon) . Thermo-couples (4 units in the PC deck) . Communication relay (Optic fiber) . Power supply
3	Bottom of the South Pylon	. GPS measurement unit (1 unit)
4	Center of the deck	. GPS measurement unit (2 units) . Anemometer (1 unit) . Air temperature Indicator (1 unit) . Rain gauge (1 unit) . Thermal-couples (4 units in the steel deck) . Thermal-couples (1 unit in the Dummy cable) . Communication relay (Optic fiber)

5	Top of the North Pylon	. GPS measurement unit (1 unit) . Monitoring camera (2 units) . Communication relay (Optic fiber) . Power supply
6	Bottom of the North Pylon	. GPS measurement unit (1 unit) . GPS station unit (1 unit)
7	Additional requested devices (Requested by MOT in April 2010)	. GPS units on the PC deck (2 units at the fouth of deck) . Thermo-couples at junction (4 units) . Strain gauges (8 units) . Fixed accelerometers (8 units)

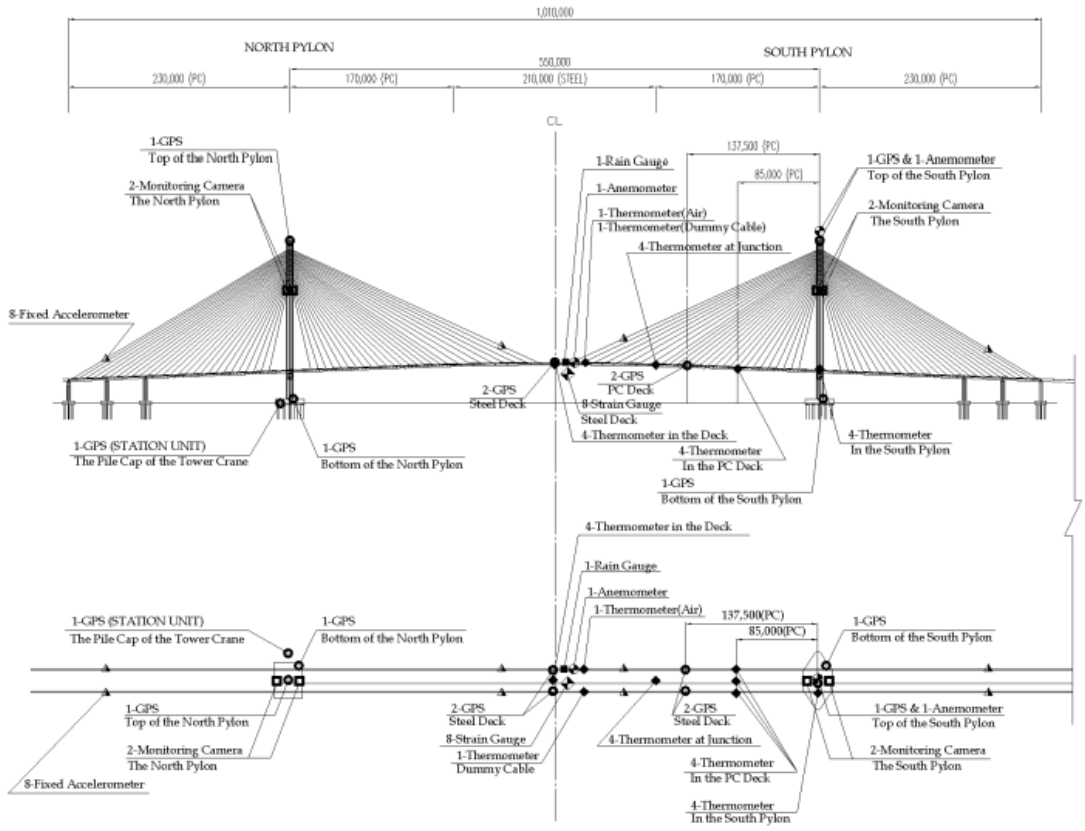


Figure 2.2 Sensor arrangement

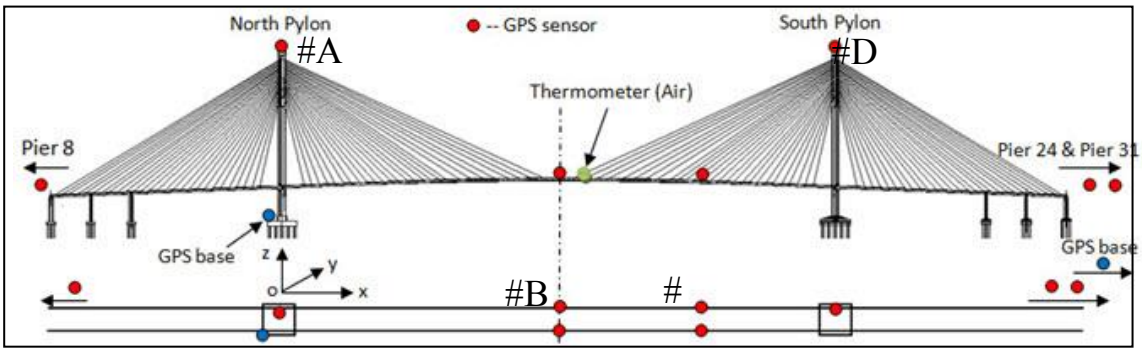


Figure 2.3 GPS sensor locations on the target bridge

2.2. GPS time-series data of the target bridge

The GPS time-series data used in this study are almost one year that were acquired from January - to November - in 2013. The frequency of the GPS system is 20 Hz. The GPS time-series data used in this study are the ten-minute average time-series data. Figure 2.4 (a)-(c) show the plots of acquired GPS time-series data (extract one week, from February 15th to February 22nd) at the center span in the x (longitudinal), y (lateral), and z (vertical) directions. The mean values of each one week data are also indicated in red line in the figure. The range of displacement around the mean coordinate in z direction, which is approximately ±0.17m, shows much larger than those in the x and y directions, while they show almost the same range order, around ±0.04m. Additionally, Figure 2.4-(d) shows the air temperature data that acquired from a thermometer placed at the center span in the same period.

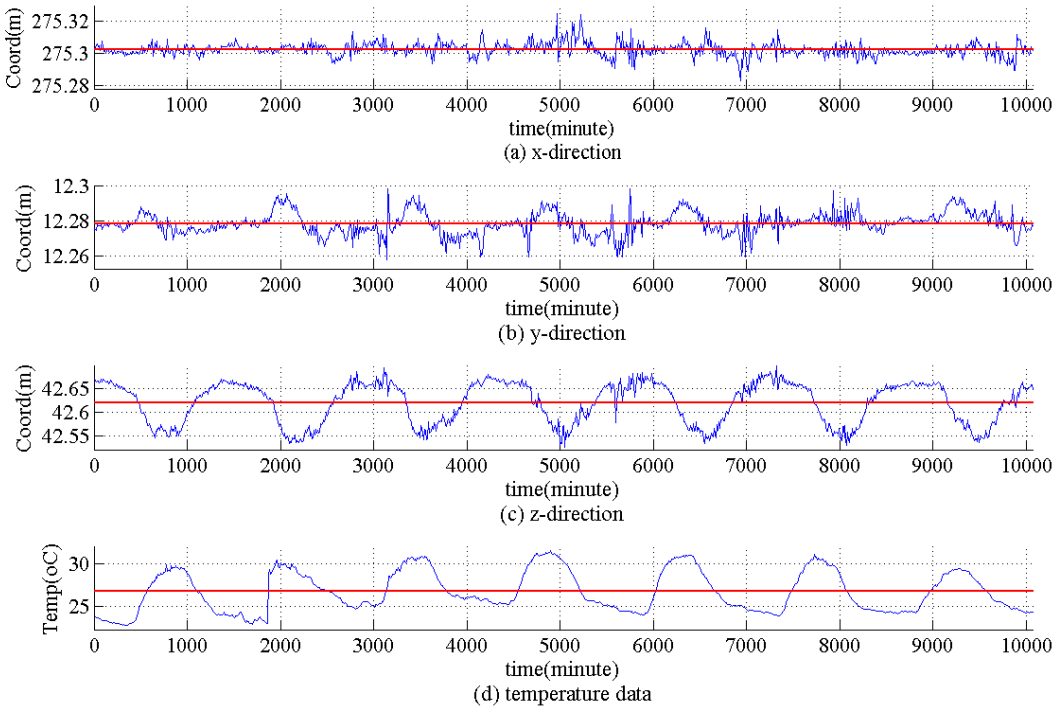


Figure 2.4. Acquired GPS data and the temperature data (blue line: observed one-week data, red line: the mean value of the one-week data)

In study of the GPS time-series observed data, there some abnormal data as well as missing data which occurred during the data acquisition. In the case of missing parts, they could be single missing data, double missing data or multiple missing data, and they occurred in many times of observation. Consequently, to interpolate the missing parts is discussed in the present study.

2.3. Handling of missing parts

There are many reasons that affect to cause the missing data during the monitoring performance such as the problem of communication system, lost record, a mistake, etc, which cannot be rectified until later. When one or more observation is missing, it may be necessary to estimate the model and also to obtain the interpolation of missing values. Actually, there are many small missing parts in the GPS time-series data of the target bridge which need to deal with before analyzing further. Statistically, almost 90% missing parts are consist of less than five missing data, were observed in the GPS time-series data. For handling those missing parts, a simple

interpolation procedure was adopted for the time-series analysis. There were actually some previous studies that verified the handling methods of missing data [21,22]. When a few data points are missing, it may be possible to interpolate the missing values by the polynomial function estimation on the basis of the least-square method [22]. Here, considering a polynomial function with *m*-th order:

$$x_i = a_0 + a_1 t_i + a_2 t_i^2 + ... + a_m t_i^m \tag{2.1}$$

For *n* data points: $(x_1, t_1); (x_2, t_2); ...; (x_n, t_n)$. Substituting into equation (2.1) will get *n* equations as shown:

$$\left.\begin{array}{c} x_1 = a_0 + a_1 t_1 + a_2 t_1^2 + ... + a_m t_1^m \\ x_2 = a_0 + a_1 t_2 + a_2 t_2^2 + ... + a_m t_2^m \\ .... \quad .... \quad .... \quad .... \\ x_n = a_0 + a_1 t_n + a_2 t_n^2 + ... + a_m t_n^m \end{array}\right\} \tag{2.2}$$

To define the coefficients $a_i (i = 0, 1, \ldots, m)$, we need at least $(m + 1)$ equations like equation (2.2). For application of Least Square Approximation, the number of equations must be larger than the number of coefficients that means $n > m + 1$.

Equation group (2.2) can be rewrite by matrix form as follows:

$$\mathbf{x} = \mathbf{M}.\mathbf{a} \tag{2.3}$$

where:

$$\mathbf{M} = \begin{pmatrix} 1 & t_1 & ... & t_1^m \\ 1 & t_2 & ... & t_2^m \\ ... & ... & ... & ... \\ 1 & t_n & ... & t_n^m \end{pmatrix}; \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix}; \qquad \mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ ... \\ a_m \end{pmatrix}$$

Hence, the coefficients of polynomial can be determined by following:

$$\mathbf{a} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{x} \tag{2.4}$$

The performance of interpolation was then verified for interpolating the missing points in acquired GPS data. A part of acquired time-series without any missing points, which was the data at the top of one of the towers in the vertical *z*-direction, was taken for the verification, and some missing points; the cases of one to five missing points, were given. The least-square interpolation

was then applied to examine the accuracy in each case. Notice that, when the number of missing data point were one or two, the difference between the interpolated and original values became less than 3mm, which was much smaller than the accuracy of the GPS system. In this example, three cases of missing data were tested to estimate that are three missing data, four missing data, and five missing data, respectively (Table 2.2). In order to apply the least square method, the order of the polynomial *m* must be determined; the accuracies of the interpolated points increase when the higher-order polynomials are used against the number of missing points. Here, to find out the simple and automatic procedure to interpolate many missing parts occurred in the raw GPS data, most of which are less than five missing data points as mentioned in the previous section. Thus, the third order was chosen as the acceptable order to interpolate as many missing parts as possible with acceptable accuracies compared to the measurement accuracy of the installed GPS system.

Table 2.2 Given data points of example

No	Actual value	3 missing case	4 missing case	5 missing case
1	168.1715			
2	168.1705			
3	168.1725			
4	168.1705	missing	missing	missing
5	168.1645	missing	missing	missing
6	168.1625	missing	missing	missing
7	168.1605		missing	missing
8	168.1665			missing
9	168.1715			
10	168.1605			

Figure 2.5 overlay the plots of the interpolated time-series and the original one in the cases of three, four, and five missing points. The residuals at the missing points are summarized in Table 2.3. It can be seen that, the accuracies of interpolated values get lower as the number of missing point increases. In the case of five missing data points, some of the absolute residuals between the interpolated and original time-series show more than 10mm, which is the 50% of the specific measurement accuracy of installed GPS system in the vertical direction. From the results of the same verification using several time-histories, it was decided that the three missing points or less were able to appropriately interpolated by the least-square based method in the target GPS

data. In addition, it was recognized that the polynomial with the order m = 3 using three previous and two posterior data points n=5 was generally able to obtain accuracies at the interpolated data points, at least less than the measurement accuracy of the GPS system. We thus applied the automatic interpolation process to many missing parts, which were with three missing points or less, in whole acquired GPS data as the pre-processing procedure for the next time-series analysis. The algorithm of the least-square method as mentioned above was used to write the codes on the MATLAB software to interpolate the missing parts automatically.

Table 2.3 Difference between the measured values and the estimated value (unit: mm)

Sample #	3 missing	4 missing	5 missing
4	+1.6	-1.4	-4.5
5	-3.4	-7.8	-13.4
6	-4.5	-9.7	-17.5
7	--	-11.0	-19.8
8	--	--	-11.3

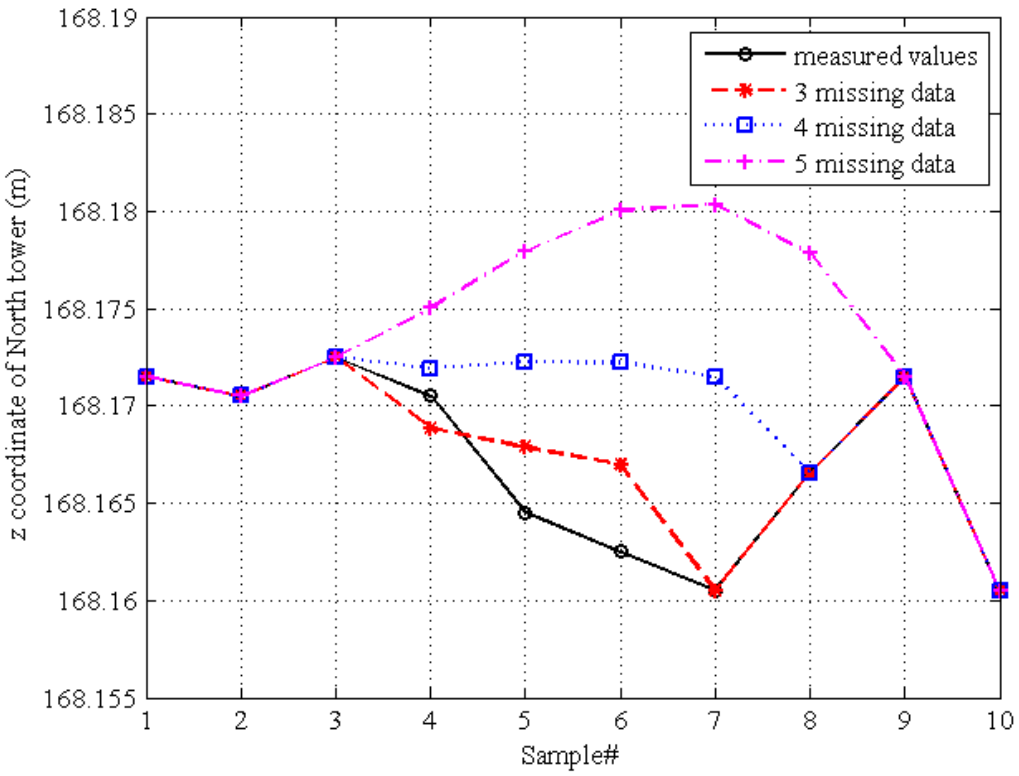


Figure 2.5 Application of least-square approximation to missing data

2.4. Correlation analysis and Global deformation

The GPS time-series data used in the time-series analysis for verification here were the pre-processed data acquired from February 15th to May 15th, 2013. In Figure 2.4, not only the GPS time-series data in (a)-(c), but also the air temperature data in the same period of observation acquired by using a thermometer placed on the center span is also shown in (d). It can be seen that there is a same periodic behavior (1440 minutes = 24 hours) between the air temperature change and the displacement from GPS especially in the *z*-direction. It was considered that the daily air temperature changes influenced on the global bridge deflection. Therefore, the correlation between the air temperature and the GPS data were required to be assessed for understanding of the global deformation mode of the target bridge under the temperature changes.

The correlation coefficient analysis was then conducted. The correlation coefficient between two variables *X* and *Y* is their covariance normalized by their standard deviations, as a following function:

$$r_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \tag{2.5}$$

Where: μ_X and μ_Y are the mean values; σ_X and σ_Y are the standard deviations of X and Y, respectively, and $E[.]$ is the expected value operator.

Four GPS locations were selected to be analyzed: the top of the north tower: #A, the top of the south tower: #D, the middle of the center span girder: #B, and the quarter of the center span girder: #C, as indicated in Figure 2.3;these locations are the typical positions to verify the behaviors of cable-stayed bridges. We then calculated the correlation coefficients between the air temperature and each of the GPS time-series data with all directions at location #A-D, and those between two of the GPS data as summarized in Table 2.4.

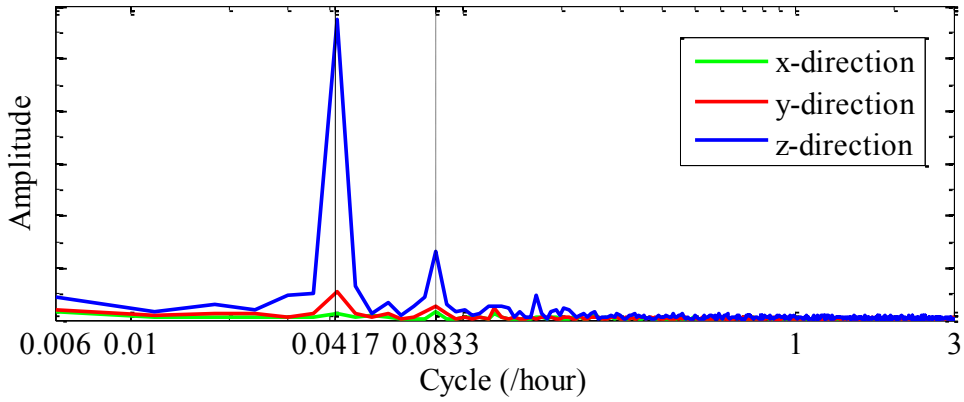
Table 2.4 Correlation coefficients between the GPS data and the air temperature data

Data		Temp.	North tower #A			Center span #B			Quarter span #C			South tower #D		
			x	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*	*x*	*y*	*z*
North	*x*	0.46	1	0.07	-0.02	0.54	0.52	-0.67	0.32	0.22	-0.56	-0.51	-0.05	0.10

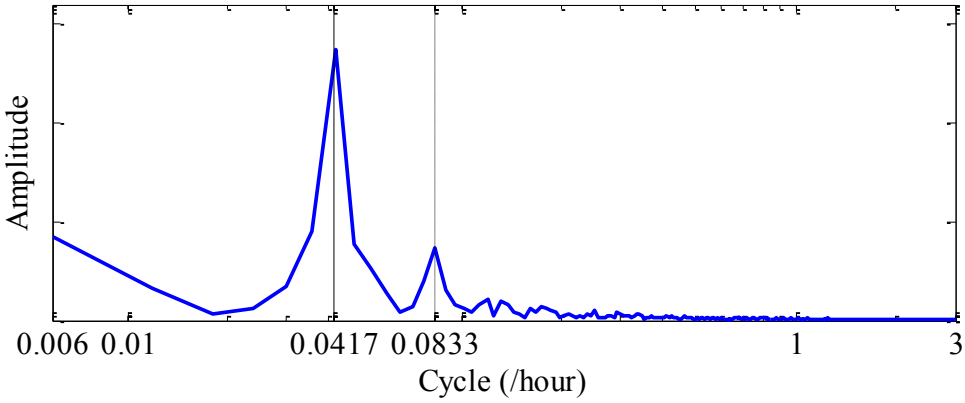
tower #A	*y*	-0.42		1	-0.36	-0.08	0.65	0.29	-0.42	0.71	0.30	-0.01	0.89	-0.33
	z	0.41			1	0.05	-0.05	-0.18	0.35	-0.09	-0.13	-0.16	-0.26	0.81
Center span #B	*x*	0.14				1	-0.02	-0.15	0.49	-0.09	-0.14	0.24	-0.13	0.08
	y	0.06					1	-0.28	-0.17	0.80	-0.21	-0.50	0.67	-0.10
	z	-0.91						1	-0.57	-0.03	0.93	0.77	0.36	-0.31
Quarter span #C	*x*	0.69			Sym.				1	-0.29	-0.58	-0.15	-0.49	0.48
	y	-0.11								1	-0.00	-0.27	0.75	-0.17
	z	-0.87									1	0.68	0.36	-0.25
South tower #D	*x*	-0.61										1	0.01	-0.22
	y	-0.49											1	-0.37
	z	0.56												1

Firstly, in seeing the correlation coefficients between the temperature and each of the GPS data at the two tower points (#A and #D), the positive and negative correlations are recognized only in the *x*-direction, respectively. It indicates that the towers deform on the inward side when the air temperature increases; therefore, the correlation coefficients in the *y*- and *z*-directions show the negative and positive values respectively in both two towers. The influences of the air temperature on the deformations of two towers are considered to be the same because those absolute values show almost the same in all directions; from 0.4 to 0.6.On the other hand, in seeing the correlation coefficients between the temperature and the GPS data from the girder points (#B and #C), it was firstly recognized that the values in the *z*-direction was much closer to -1; i.e., the negative correlation, both in #B and #C. This clearly indicates that the girder shows the global deflections as the air temperature increases. Moreover, the other correlation coefficients in the *x*- and *y*-directions show low correlations except the one in the *x*-direction at the quarter point of the girder (#C). This was considered to be because the target cable-stayed bridge was almost symmetric; therefore, the displacement in the *x*-direction; i.e., the longitudinal direction, did not occur especially at the center span (#B). The high negative correlation at the quarter span (#C) in the *x*-direction was then observed because the location there was the non-symmetric position of the girder. Lastly, the correlation coefficients in the *y*-direction at both #B and #C are very small; therefore, it can be said that the in-plane deformation (the *x*-*z* coordinate) of the cable-stayed bridge is dominated in the global deformation of the girder.

In addition, the frequency (/hour) spectra of the GPS data and their temperature data were derived to compare the dominated frequencies of those time-series. Figure 2.7 (a) and (b) are obtained spectra; (a) is from the one-week GPS data at the center span, which is exactly the time-series presented in Fig.3 (a)-(c), and Figure 2.7 (b) is from the temperature data in Figure 2.4 (d). Here, the dominated frequency (/hour) components in the target time-series can be examined within the range from 0.00595 /hour (=1week, 7days) to 3 /hour (20 minutes). It can be seen that the same two frequencies show high amplitudes; the highest one is 0.0417 /hour (= 24 hours) and the other is 0.0833 /hour (= 12 hours) in two spectra: the GPS data in the *z*-direction and the air temperature data. Notice that the GPS data in the *z*-direction at the center span show the highest absolute correlation coefficient. This indicates that the GPS data show high correlation with the air temperature, are mostly dominated by the one-day periodic behavior, which is the 24 hours-cycle; the highest and lowest temperatures are thus observed at the noon and at the midnight, respectively.



(a) GPS data at the center span



(b) Air-temperature data

Figure 2.7 Comparison of the spectra of the GPS data and the temperature data

On the basis of those results, the global deformation modes due to the periodic temperature changes could be identified. Here, the one certain day of GPS time-series data were extracted to validate the global deformation behaviors as shown in the Figure 2.8. Then, the *x*- and *z*- displacements at the midnight and the noon calculated by differencing the *x*- and *z*-coordinates from those at 6am in a certain day were plotted in the Figure 2.9 with four interested points: #A-D. Notice that the same symbols show the displacements at the same time of observation. It can be understood that the global deformation mode due to the increase of air temperature is configured by the inward deformations of the two towers and the vertical deflection of the girder, and the reverse occurs when the temperature decreases. In seeing the correlation coefficients between two of GPS data presented in Table 2.4, high absolute correlations can be observed not only between the *z*-directions at the two points of the girder, but also between two of all directions at the two towers, and between the *z*-direction at the girder and the *x*-direction at each of the towers.

From those results, the global deformation due the air temperature changes was recognized. It was then concluded that the displacements of the two towers and the girder in the deformation due to the one-day periodic temperature change are highly correlated; therefore, the statistical pattern recognition of those global deformation was expected to be applicable for assessing the condition of the structure globally using the GPS data.

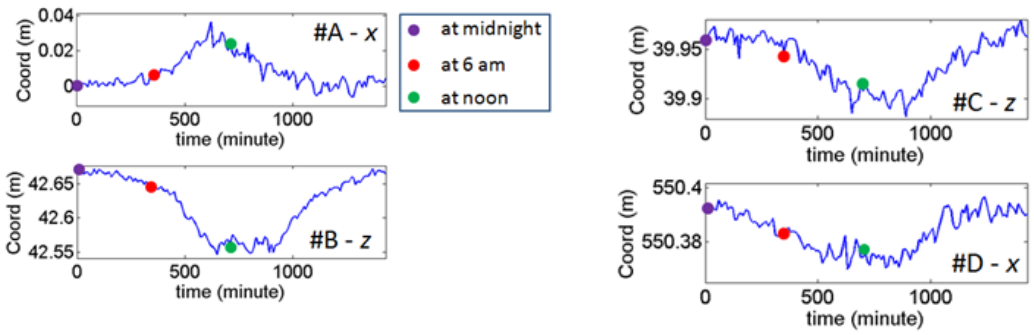


Figure 2.8 One certain day of time-series data

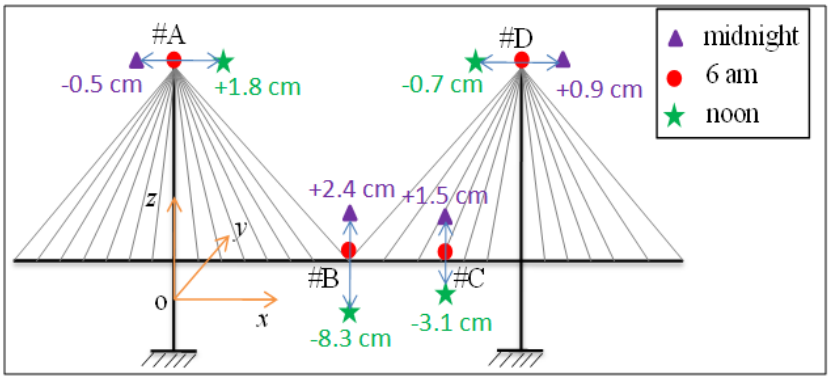


Figure 2.9 In-plane global deformation modes due to the day-cycle periodic temperature change

2.5. Summary

The Chapter 2 presents the target bridge of current research and the GPS observed data studying is also shown. The target bridge is generally presented in section 2.1 as well as the GPS system used in the SHM of the target bridge. Additionally, one week GPS time-series data in ten minute average of center span point are extracted to show in section 2.2, whereas the displacements in *z*-direction shows much larger than those in *x* and *y*-directions. To interpolate the missing data parts in time-series data, the least square approximation method is present to apply in section 2.3. The algorithm of the least square method was used to write codes on MATLAB software to automatically interpolate the missing parts of the GPS time-series data. Furthermore, the correlation between time-series data and the air temperature data as well as the frequency spectra of them are analyzed in section 2.4. The results of the correlation analysis show understanding of the global deformation of the target bridge.

Chapter 3

Autoregressive Integrated Moving Average model and its application in GPS time-series data

On the basis of the considerations in the previous chapter, we then verified the applicability of the time-series analysis to the feature extraction for the structural condition assessment that statistically investigated the pattern of the global deformations. Here, the autoregressive integrated moving average (ARIMA) model was adopted because the global deformation was understood to be dominated by the one-day periodic behavior. From the results in the previous chapter, we then picked up some data to be analyzed; they were the data in all of three directions at the top of two towers, the data in the *z*-direction at the center span of the girder, and the data in the *x*- and *z*-directions at the quarter span, all of which showed the high correlation coefficients with the air temperature data.

3.1. Theory of time series analysis

3.1.1. Stationary and non-stationary process

A time series is called as a stationary series if its behavior does not change over time, and the stationarity is a key idea in time-series analysis. This means that the values always tend to vary about the same level and that their variability is constant over time [23]. In other words, we can look at a stationary time series from an intuitive point of view. If the property of one section of the data is much like those of any other section, then we can call the "stationary" series. This means that the stationary series has no systematic change in mean and variance. In addition, any periodic variations must be removed. Figure 3.1 shows an example about the flow volume of the Nile at Aswan from 1871 to 1970 [23]. The general pattern of this data does not change over time by an intuitive point of view, so it can be seen as a stationary series. The theory of stationary series is well known and it is very popular that mentioned in various studies. This means that they play a fundamental role in the study of time series. Obviously, not all time series are stationary.

Indeed, non-stationary series tend to be the rule rather than the exception. However, many time series are related in simple ways to the stationary series [23].

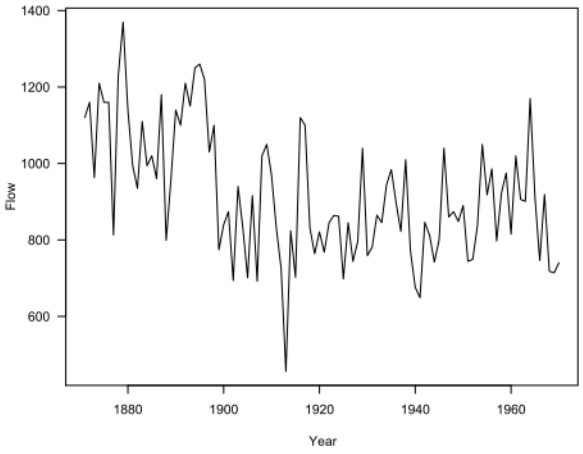


Figure 3.1 Flow volume of the Nile at Aswan from 1871 to 1970 [23]

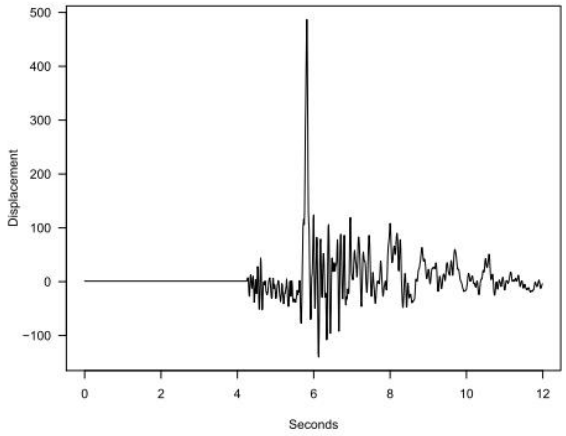


Figure 3.2 Horizontal ground displacement in an earthquake [23]

A time series process is non-stationary if it appears to have no fixed level, and the time series may also display some periodic fluctuations. There is an example about ground displacement in an earthquake that is shown in Figure 3.2 that as an example in non-stationary series. Practically, most of the time series tools only apply to stationary time series. In other words, we have to examine the data and try to transform any non-stationary data set into a stationary model. There are two main types of transformations to make the series stationary. Firstly, the time series data can be transformed through differencing if the stochastic process has an unstable mean. This type of transformation is used for the purpose of removing the polynomial trend that is exhibited by the data. The logarithmic and square root transformations are special cases of the class of transformations. These transformations are used if the series being examined has a non-constant mean and variance and it results in a straighter curve plot [24, 25].

3.1.2. Autoregressive Moving Average models (ARMA models)

a) Autoregressive models (AR models)

The basic AR model is a model that contains only one AR parameter as the series follows:

$$X_t = \phi X_{t-1} + \varepsilon_t \tag{3.1}$$

Where: $\{\varepsilon_t\}$ is a normal random variable or white noise with mean zero and variance σ_z^2. It means that any given value X_t in the time series is directly proportional to the previous value X_{t-1} plus some random error ε_t. As the number of AR parameters increase, the series X_t becomes directly related additional past values. The general autoregressive model with p AR parameter can be written as:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \varepsilon_t \tag{3.2}$$

Where: $\phi_i (i = 1 \div p)$ are the AR parameters; the p-value is called as the number of orders of the AR models. Hence, it is called an autoregressive model of order p and it is usually abbreviated as AR(p). If we call B is the backward shift operator such that:

$$B^j x_t = x_{t-j} \tag{3.3}$$

and:

$$\phi(B) = 1 - \phi_1(B) - \ldots - \phi_p(B^p) \tag{3.4}$$

The values of ϕ which make the process stationary are such that the roots of $\phi(B) = 0$ lie outside the unit circle in the complex plane [23,24].

b) Moving Average models (MA models)

Unlike the autoregressive model, the Moving Average model parameters relate what happens in period t only to the random errors that occurred in past time periods. The basic MA model with one parameter can be written as follows:

$$X_t = \varepsilon_t - \theta_t \varepsilon_{t-1} \tag{3.5}$$

It means that any given X_t in the time series is directly proportional only to the random error ε_{t-1} from the previous period plus some current random ε_t. The general moving average with q MA parameters can be written as:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \tag{3.6}$$

Where $\theta_j (j = 1 \div q)$ are the MA parameters; the q-value is called as the number of orders of the MA models. Hence, it is called a moving average model of order q and it is usually abbreviated as MA(q) [24,25].

c) Autoregressive Moving Average models (ARMA models)

In most cases, it is the best way to develop a mixed autoregressive moving average model when building a stochastic model to represent a stationary time series. The order of an ARMA model is expressed in terms of both *p* and *q*. The model parameters relate to what happens in period *t* to both the past values and the random errors that occurred in past time periods.

A general ARMA model can be written as follows:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... - \theta_q \varepsilon_{t-q} \quad (3.7)$$

By using a backward shift operator *B* as mentioned above:

$$B^j x_t = x_{t-j}$$

$$\phi(B) = 1 - \phi_1(B) - ... - \phi_p(B^p) \quad (3.8)$$

(3.9)

$$\theta(B) = 1 - \theta_1(B) - ... - \theta_q(B^q) \quad (3.10)$$

The equation (3.7) can be simplified to obtain as follows:

$$\phi(B) X_t = \theta(B) \varepsilon_t \quad (3.11)$$

3.1.3. Autoregressive Integrated Moving Average models (ARIMA models)

In general, most of the time series are non-stationary, and as mentioned in the section 3.1.3, they need to remove any non-stationary sources of variation, and then fit a stationary model to the time series data. Practically, we can achieve stationarity by applying regular differences to the original time series. In order to solve the non-stationary process, the ARIMA model is one of statistical models to describe non-stationary time-series.

In definition of an ARIMA model [24], a process X_t is called to be ARIMA (*p,d,q*) if:

$$\nabla^d X_t = (1 - B)^d X_{t-1} \quad (3.12)$$

is ARMA(*p,q*) model. Where: ∇ is differencing operator.

In general, the model can be written as:

$$(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p)(1-B)^d X_t = c+(1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q)\varepsilon_t \quad (3.13)$$

Or:

$$\phi(B)(1-B)^d X_t = c+\theta(B)\varepsilon_t \quad (3.14)$$

Again, here: $\phi(B)=1-\phi_1(B)-\ldots-\phi_p(B^p)$

and: $\theta(B)=1-\theta_1(B)-\ldots-\theta_q(B^q)$

Overall, an ARIMA(*p,d,q*) model contains three terms that are autoregressive term, integrated term, and moving average term with the p, d, and q orders respectively. In each case of applying the ARIMA model, which one is the most appropriate model that can be defined by two ways; first of all is the application of autocorrelation function (ACF) and partial autocorrelation function, and other one is the Bayesian information criterion method. In some case of analysis, it can be combined both two ways to define the order values.

a) Autocorrelation function and partial autocorrelation function.

The Autocorrelation function (ACF) is one of the most important tools in the identification stage for building time series models. It measures how strongly time series values at a specified number of periods apart are correlated to each other over time. The number of periods apart is usually called the lags. For instance, an autocorrelation for lag 1 is a measure of how successive values are correlated to each other throughout the series. An autocorrelation for lag 2 measures how series values that are two periods away from each other are correlated throughout the series.

The autocorrelation values may range from negative one to positive one, where a value close to positive one that indicates a strong positive correlation, a value close to negative one that indicates a strong negative correlation, and a value close to zero indicates no correlation.

The autocorrelation function between X_t and X_{t-m} is defined as:

$$\rho_i = corr(X_t, X_{t-m}) = \frac{\mathrm{cov}(X_t, X_{t-m})}{\sqrt{\mathrm{var}(X_t)\,\mathrm{var}(X_{t-m})}} \quad (3.15)$$

The partial autocorrelation function (PACF) is another important tool that similar to autocorrelation function. By PACF, a set of partial autocorrelation values would be computed for a given time series corresponding to a range of lag values which are used to evaluate relationships

between time series values. Partial autocorrelation values also range from negative one to positive one, and are very useful in some situations where the autocorrelation patterns are hard to determine. In practice, the estimated autocorrelation function is used in place of the theoretical autocorrelation function to generate estimates of the partial autocorrelation function.

In case of system identification, for example, given a set of observation $X_1, X_2, \ldots, X_t$, we need to decide what the appropriate model might be. The estimated ACF and PACF are the tools which can be used to do this. If the ACF exhibits slow decay and the PACF cuts off sharply after lag p, we would identify the series as AR(p). If the PACF shows slow decay and the ACF show a sharp cutoff after lag q, we would identify the series as being MA(q). If both the ACF and PACF show slow decay we would identify the series as being mixed ARMA. In this case, the orders of the AR and MA parts are not clear, but it is reasonable to first try ARMA(1,1) and move on to higher order models if the fit of this model is not good [23].

b) Bayesian information criterion (BIC) method

Practically, when the ACF and PACF do not show clear in the order of the AR and MA, another method can be applied to decide which model is appropriate that is called Bayesian information criterion (BIC) method. The BIC method is also called the Schwarz Information Criterion (SIC) for an approach yielding the same statistic based on a minimum description length argument. The definition of BIC can be shown as [24,28]:

$$BIC = \log \hat{\sigma}_k^2 + \frac{k.\log n}{n} \tag{3.16}$$

Where: $\hat{\sigma}_k^2$ is given by:

$$\hat{\sigma}_k^2 = \frac{SSE_k}{n} \tag{3.17}$$

SSE_k denotes the residual sum of squares under the model with k regression coefficients.

k is the number of parameters in the model.

n is the sample size.

The key of this method in application is that the value of k yielding the minimum BIC specifies the best model. Various simulation studies have tended to verify that BIC does well at getting the correct order in large samples.

c) Augmented Dickey-Fuller Unit Root test

In statistic, an Augmented Dickey-Fuller test (ADF test) is a test for a unit root in a time series sample. In other words, this kind of test answers us the question that how do we know when to difference time series data to make a stationary series [29,30]. The testing procedure for the ADF test is the same as for the Dickey-Fuller test. The unit root test of a series is carried out under the null hypothesis $H_0 = 0$ against the alternative hypothesis of $H_1 < 0$. One a value for test statistic is computed as:

$$DF_\tau = \frac{\hat{\gamma}}{SE(\hat{\gamma})} \tag{3.18}$$

where: $\hat{\gamma}$ is the least square estimate; $SE(\hat{\gamma})$ is usual standard error estimate.

The test statistic value can be compared with the relevant critical value for the Dickey-Fuller test. If the test statistic value is less than the critical value, then the null hypothesis of $H_0 = 0$ is rejected, and no unit root is present.

d) Ljung-Box Q test

The Ljung-Box Q test is also known as Ljung-Box test that is type of statistical test whether any of a group of autocorrelations of a time series is different from zero. Generally, the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) are useful qualitative tools to assess the presence of autocorrelation at individual lags. However, the Ljung-Box Q test is more qualitative way to test for autocorrelation at multiple lags jointly [31,32]. The formal definition of Ljung-Box Q test can be determined as follows:

$H_0 : \rho_1 = \rho_2 = ... = \rho_m = 0$; the null hypothesis means that the fist *m* autocorrelation are jointly zero. In this test, the selection of *m* affects test performance. Practically, if the length of a observed time series X_t is *N*, choosing $m = \ln(N)$ is recommended for testing power, and the multiple *m* can be tested generally [33,34].

The Ljung-Box Q test is given by:

$$Q(m) = N(N+2)\sum_{h=1}^{m}\frac{\hat{\rho}_h^2}{N-h} \tag{3.19}$$

Under the null hypothesis, $Q(m)$ follows the χ^2_m distribution.

The theory of Ljung-Box Q test was used to build the 'lbqtest' function on MATLAB software. The performance of Ljung-Box Q test for the testing residuals from a time series model can be followed:

- The null hypothesis: residual is white noise

- The alternative hypothesis: residual is not white noise

The null hypothesis is accepted or rejected that based on the *p*-values and the statistic values. In both cases, a *p*-value is calculated as the probability past $Q(m)$ in the relevant distribution. If the *p*-values are less than the significant level (for instance, 5%) that indicates the possibility of non-zero autocorrelation within the first m lags, meaning that the null hypothesis is rejected. Futhermore, in comparison between the statistic values and the critical values, if the calculated statistic values are less than the critical values at any lags, the null hypothesis is not rejected.

3.2. Application of ARIMA model in GPS time-series data of the target bridge

Here, the autoregressive integrated moving average (ARIMA) model is adopted because the global deformation was understood to be dominated by the one-day periodic behavior. From the results in the previous chapter, we then picked up some data to be analyzed; they were the data in all of three directions at the top of two towers, the data in the *z*-direction at the center span of the girder, and the data in the *x*- and *z*-directions at the quarter span, all of which showed the high correlation coefficients with the air temperature data.

3.2.1. ACF and PACF analysis

In this case of study, a part of the one of the GPS time-series data was extracted to use as an example, which is the one-week data (from Feb. 15th to 22nd, 2013) in the *z*-direction acquired at the center span of the girder as shown in Figure 3.3(a). In an intuitive point of view, there is periodic that appears in the GPS observed data. Moreover, this periodic is clearly shown in the its

ACF results in the Figure 3.3(b), where the apparent 144 points; equal to 1-day, periodic behaviors with a slow damping is clearly observed; it can thus be said that the data is non-stationary and influenced by the 1-day periodic temperature change.

As shown in the theory of time series analysis, the differencing method needs to apply for transforming the non-stationary series to the stationary series. Therefore, the first order of difference (d = 1) was taken to apply for the GPS observed data, then the ACF and PACF were applied again as shown in the Figure 3.3(c – e). The application of these methods were performed by using the "diff" function and "autocorr; parcorr" functions on MATLAB software.

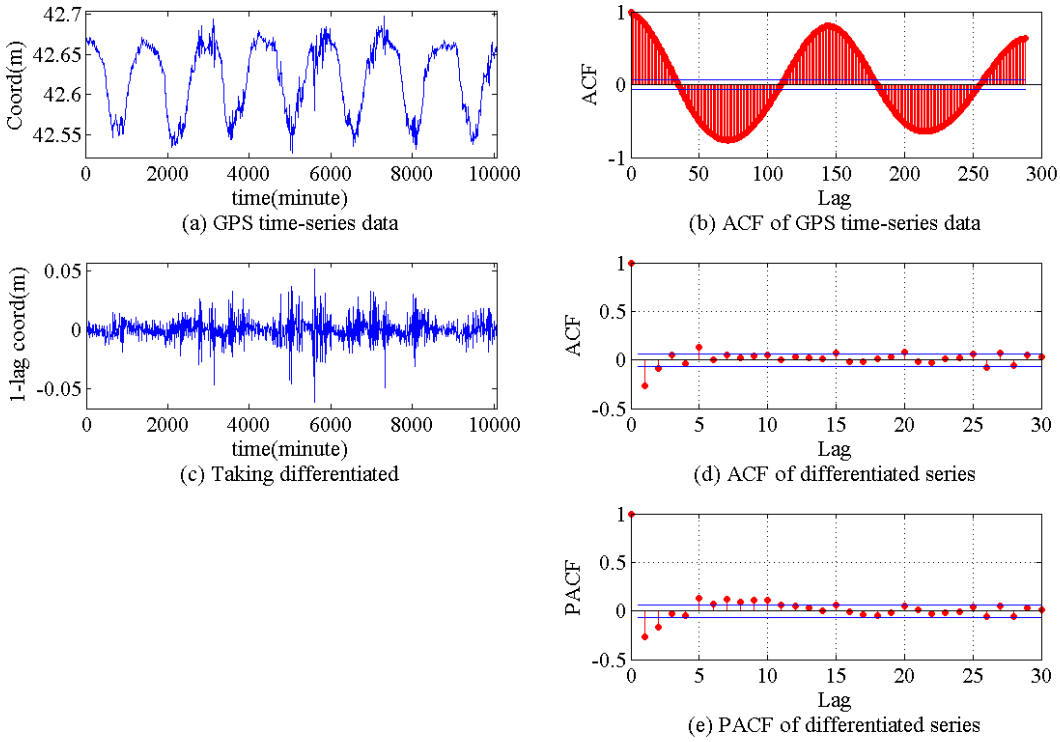


Figure 3.3 ACF and PACF of GPS time series and its differentiated time series

It can be seen that, the values of series always tend to vary about the same level and that their variability seems to be constant over time after taking difference. Besides, the periodic was already removed (Figure 3.3(c)). Furthermore, looking at the ACF and PACF results of differencing series, the ACF and PACF then dropped quickly to within the 95% confidential interval indicated by blue lines, as shown in Figure 3.3(d-e). Therefore, it can be said that the GPS observed data was transformed to stationarity.

Moreover, to check whether the differentiated time-series was stationary or not, the Augmented Dickey Fuller (ADF) test in statistical analysis can be also conducted [24,28]. In the theory of ADF test, the statistical possibility (*p*-value) and the absolute critical value are the keys to decide the null hypothesis of a test is accepted or rejected. In this current study, the ADF test was also applied for the differentiated time series to check it is stationary or not. The null hypothesis of ADF test here is "the first order differentiated data was non-stationary". This test was also performed by using “adftest” function on MATLAB. The results of this test show that, the possibility value (*p*-value) is 0.001 (0.1 %) that is much smaller than the 5% of significant level; and the absolute test statistic value is 41.51 that is much larger than the absolute critical value (3.41).Therefore, the null hypothesis here is rejected; it indicated that the original time-series was transformed to stationarity by taking the first order difference.

When the time-series became stationary by taking the difference, the orders of AR process *p* and MA process *q* in ARIMA(*p*, *d*, *q*) were determined by the partial autocorrelation function (PACF) and the ACF. Those for the differentiated time-series are shown in Figure 3.3(d) and (e). It can be seen that the values are almost dropped inside the 95%confidential interval at lag=2 in ACF and lag=3 in PACF; therefore, q =1 and p=2 can be accepted for the MA and AR orders, respectively. However, no values after the identified lag are completely inside the confidential interval. To determine those orders *p* and *q* clearly, the Bayesian information criterion (BIC) method was additionally examined.

The verifications were then conducted by calculating the BICs to the ARIMA models for several GPS time-series data using the MATLAB function "aicbic". Here, the $\log \hat{\sigma}_k^2$ in the BIC formula is the optimized log-likelihood function obtained in the estimation of each ARIMA model. The number of the order values are examined in this case that is from 1 to 4. One of the results is obtained from an one-week GPS data acquired at the center span, is shown in Table 3.1. The row the matrix corresponds to the AR order *p* from 1 to 4, and the column is the MA degree *q* also from 1 to 4. Notice that the order *d* of ARIMA(*p*, *d*, *q*) is all d=1 on the basis of the results discussed for Figure 3.3.It can be seen that the smallest BIC value is calculated in(p, q)=(1,1); the orders *p*, *d*, and *q* were thus determined to ARIMA(1, 1, 1) that is the appropriate model to apply for GPS data in this current study.

Table 3.1 BIC results for the definition of ARIMA model order

	MA(*q*)

		1	2	3	4
AR(*p*)	1	-7847	-7846	-7839	-7824
	2	-7844	-7832	-7824	-7832
	3	-7838	-7824	-7826	-7828
	4	-7833	-7826	-7829	-7823

3.2.2. ARIMA(1,1,1) model estimation

The coefficients of the ARIMA model estimated for the GPS data that consists of one-day periodic behaviors were considered to be applicable as the feature to indicate the global deformation modes of the target *Can Tho* bridge. We verified this by applying the ARIMA model estimation to the data from February 15th to May 15th, 2013. Notice that the data to be analyzed were the data in all of three directions at the top of two towers (#A and #D), the data in the *z*-direction at the center span of the girder (#B), and the data in the *x*- and *z*-directions at the quarter span (#C), as mentioned in the previous chapter.

The GPS data were taken to divide into day by day, and ARIMA(1,1,1) model was estimated to each of one-day time-series. The Figure 3.4 explains how to apply the ARIMA(1,1,1) model by taking three day examples of time-series data. This process was used for the application of ARIMA model in all GPS time-series data. Figure 3.5 shows one result that is from one of one-day GPS time-series in the *z*-direction at the center span. The time-series from the estimated ARIMA model and the original one are overlaid in Figure 3.5(a), and Figure 3.5(b) is the standardized distribution of the residual errors with the standardized normal distribution. In addition, the results of ARIMA(1,1,1) model for the first one week (7 days) of a significant direction of the target bridge were summarized from Figure 3.6 to Figure 3.12, where they are the overlaid plot between original series and fitted one, residual and its distribution plots, and the ACF, PACF of the residuals. It can be seen that, the ARIMA(1,1,1) is fitted quite well for day by day of time series data, the residual distribution plots look like the standardized distribution.

Here, the Ljung-Box Q-test (LBQ-test) [24] was also conducted to check whether the residual error distribution showed the white noise process with the normal distribution. The null hypothesis here was; "the residual was the white noise."In the results, the statistical possibility *p*-values at lags (5, 10, 15) were (0.0624, 0.1668, 0.3410) that were all larger than the 5% significant

level, and the test statistic values were calculated at (7.3199, 11.6650, 14.4788), which were smaller than the critical values at 5% significant level (7.8147, 15.5073, 22.3620). It means that the null hypothesis was not rejected. The residuals were thus identified as the white noise process. We also checked those performances in the estimations to other one-day time-series and confirmed that the estimations were conducted with almost the same and appropriate accuracies.

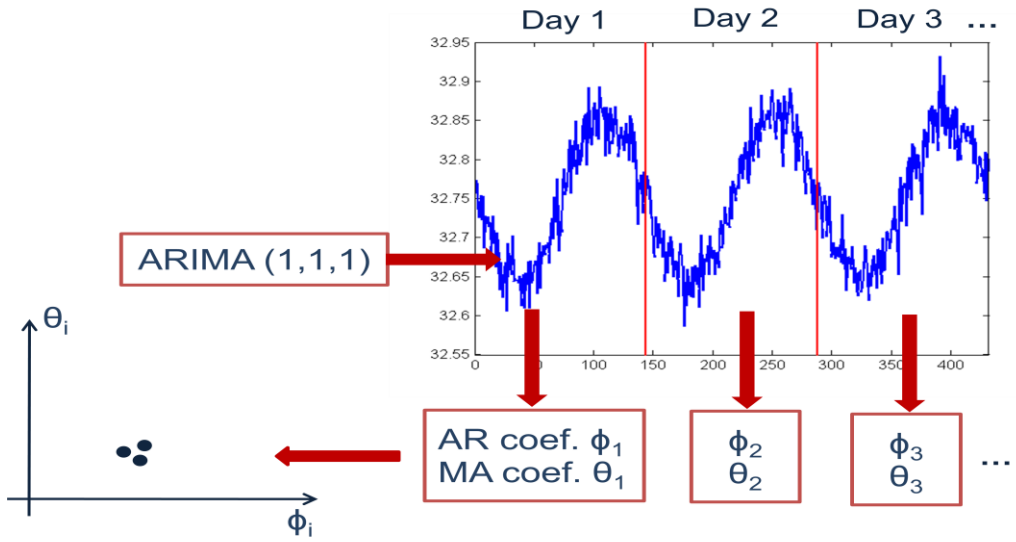
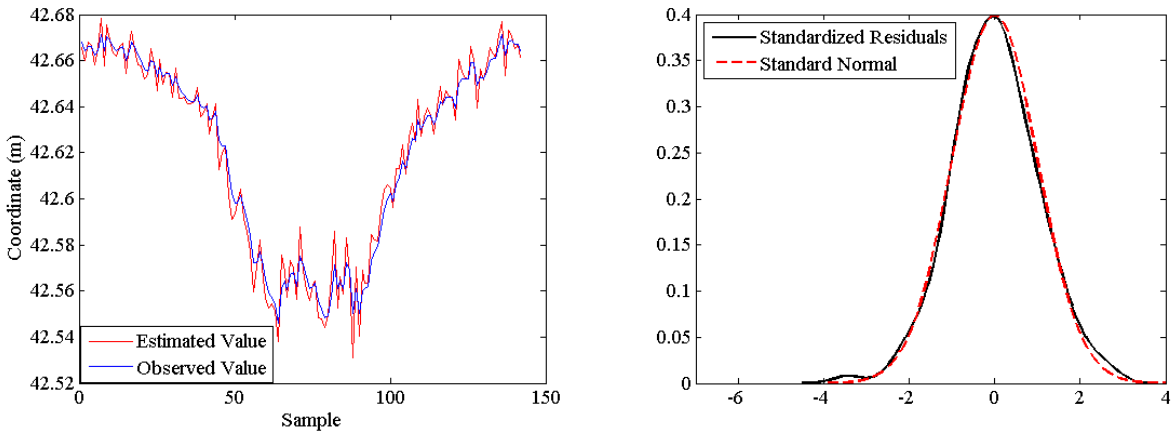


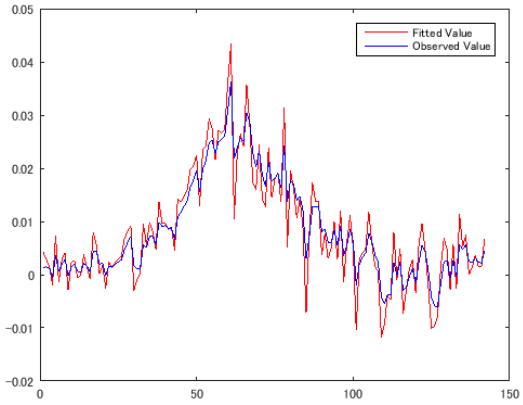
Figure 3.4 Process of ARIMA application



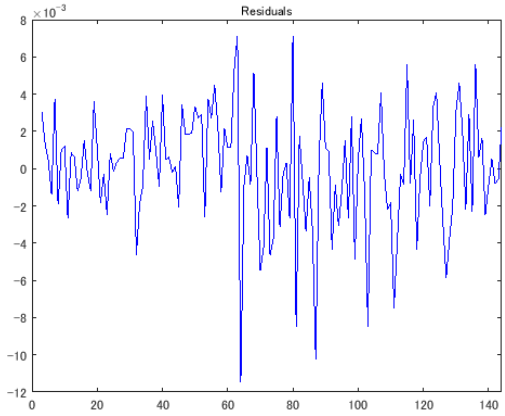
(a) Overlay of estimated and measured time-series

(b) Distribution of residual error

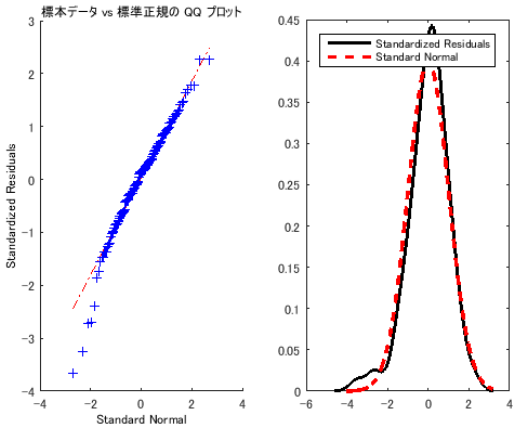
Figure 3.5 One results of estimation of ARIMA(1,1,1) model



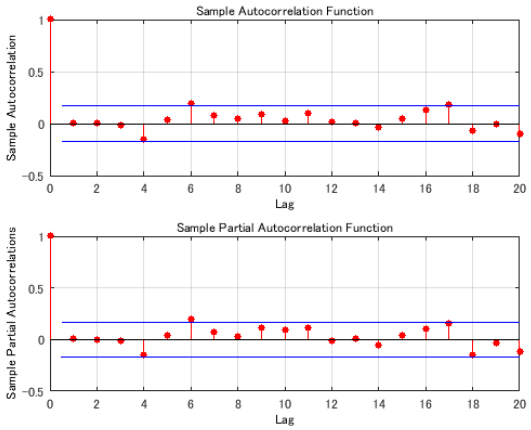
(a) Overlaid plot



(b) Residual

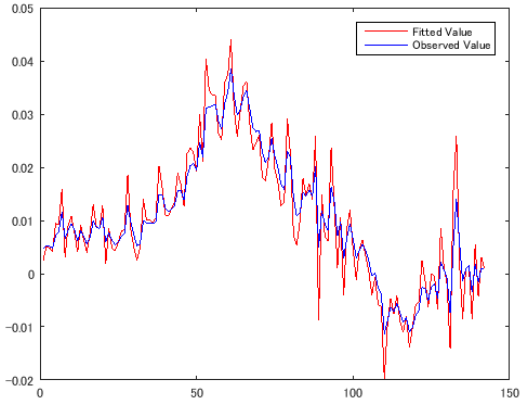


(c) Distribution of residual error

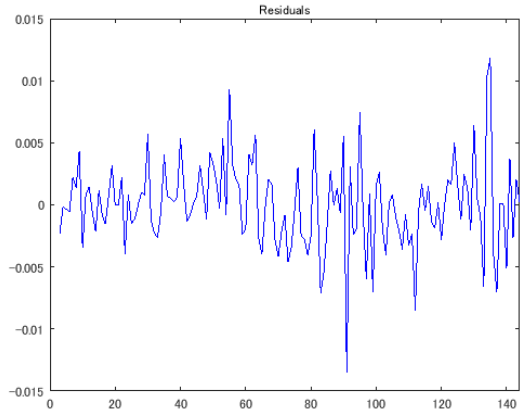


(d) ACF and PACF of residual

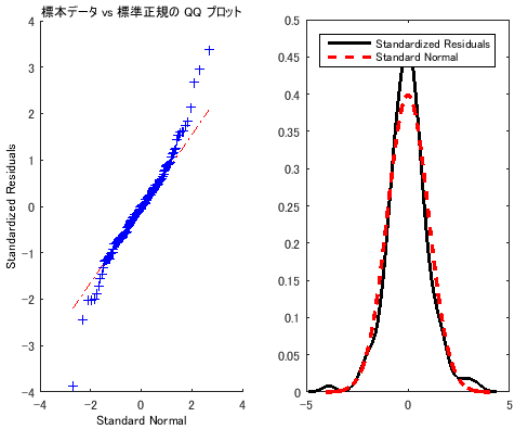
Figure 3.6 Application of ARIMA(1,1,1) model in Day 1

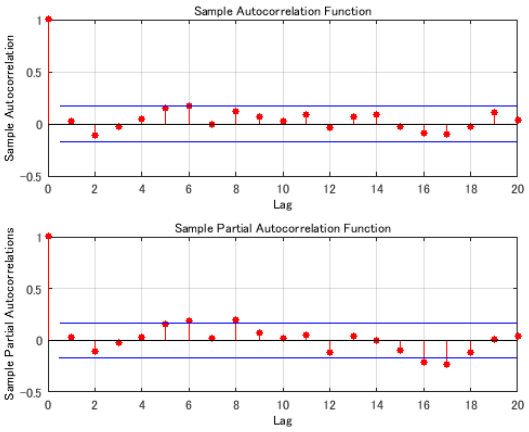


(a) Overlaid plot



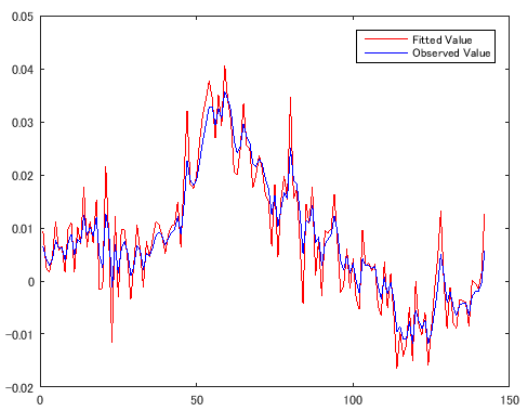
(b) Residual

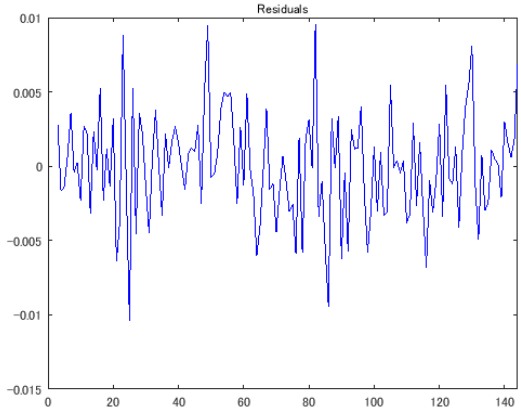




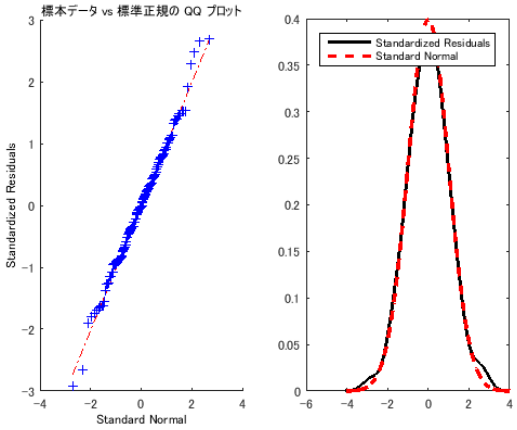
(c) Distribution of residual error (d) ACF and PACF of residual

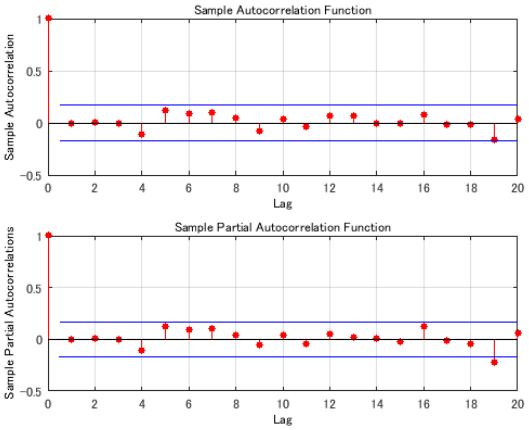
Figure 3.7 Application of ARIMA(1,1,1) model in Day 2





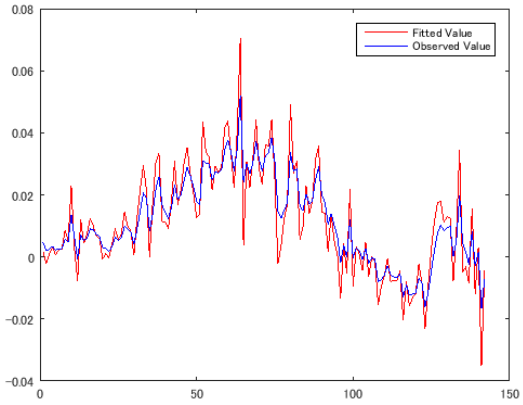
(a) Overlaid plot (b) Residual



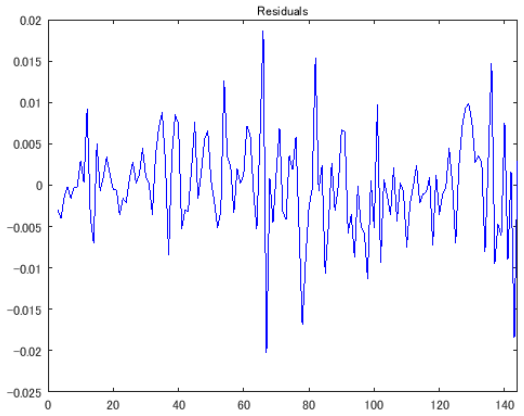


(c) Distribution of residual error (d) ACF and PACF of residual

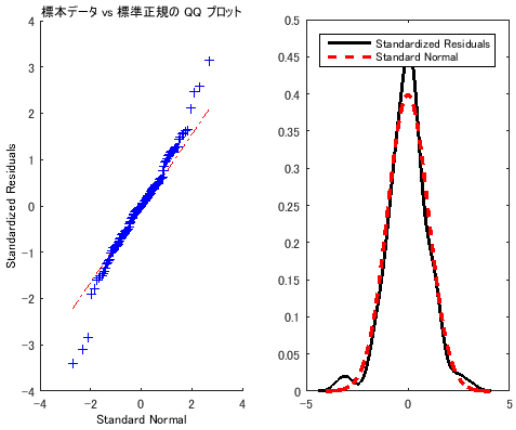
Figure 3.8 Application of ARIMA(1,1,1) model in Day 3



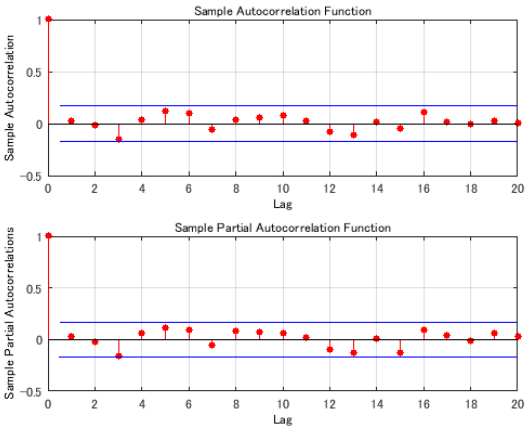
(a) Overlaid plot



(b) Residual

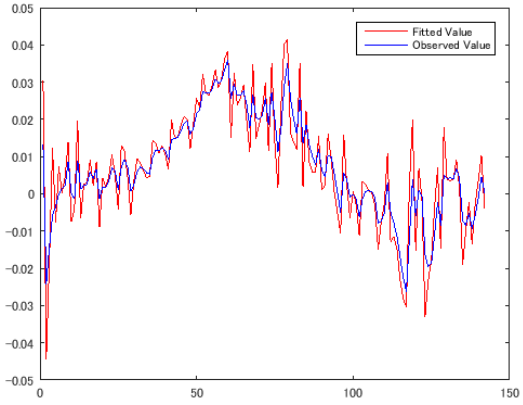


(c) Distribution of residual error

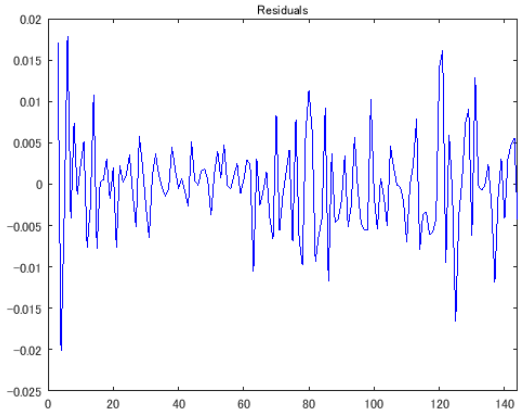


(d) ACF and PACF of residual

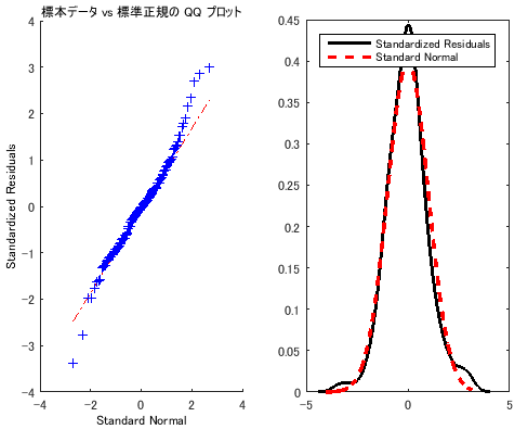
Figure 3.9 Application of ARIMA(1,1,1) model in Day 4

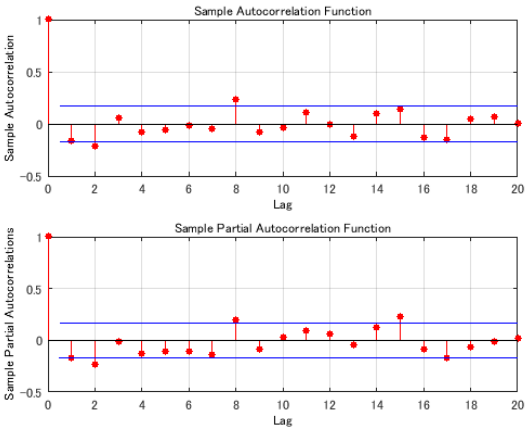


(a) Overlaid plot



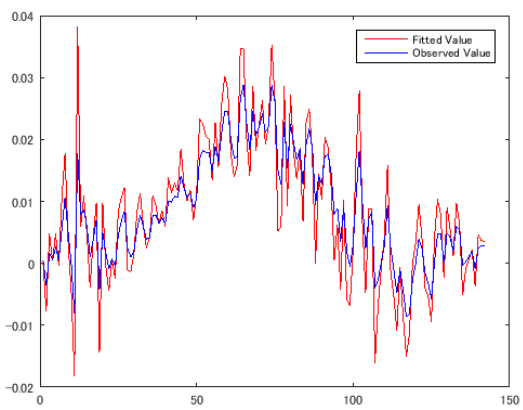
(b) Residual

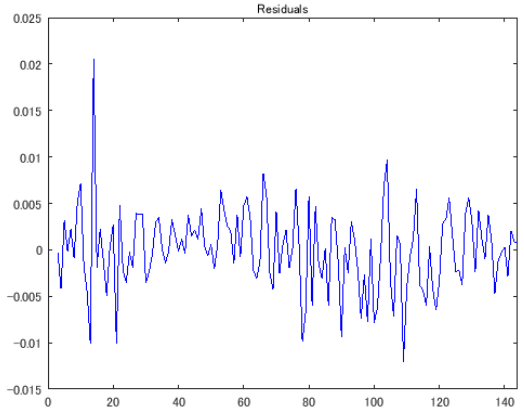




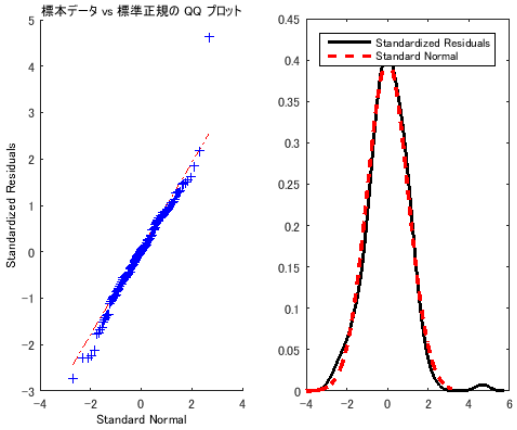
(c) Distribution of residual error (d) ACF and PACF of residual

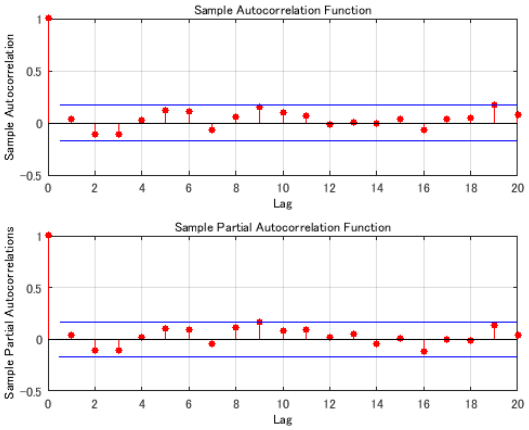
Figure 3.10 Application of ARIMA(1,1,1) model in Day 5





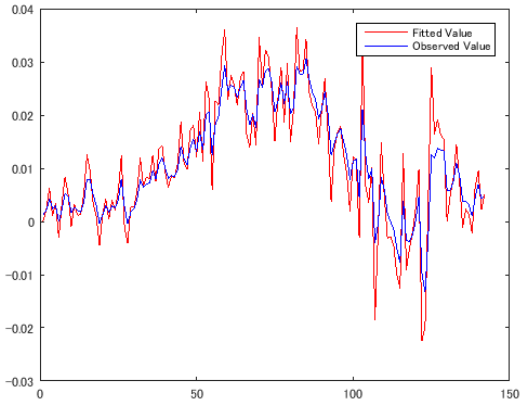
(a) Overlaid plot (b) Residual



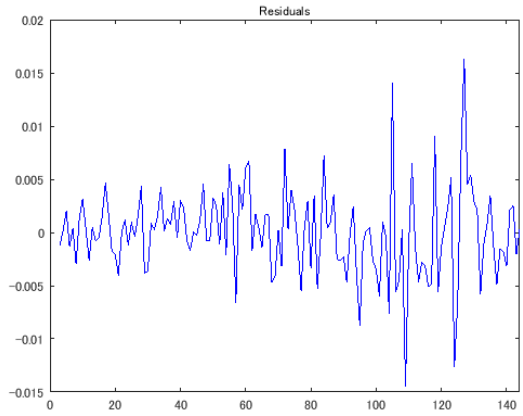


(c) Distribution of residual error (d) ACF and PACF of residual

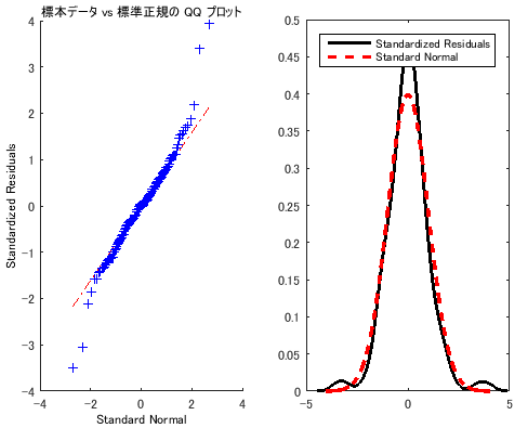
Figure 3.11 Application of ARIMA(1,1,1) model in Day 6



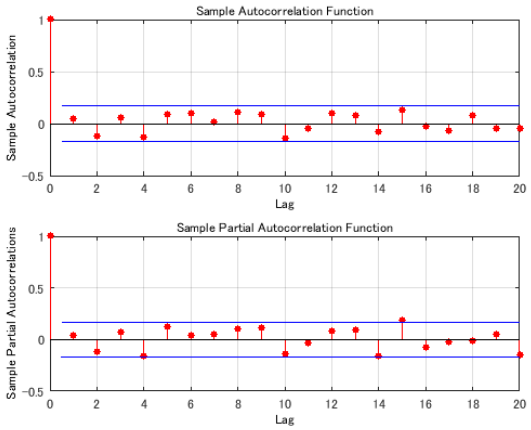
(a) Overlaid plot



(b) Residual



(c) Distribution of residual error

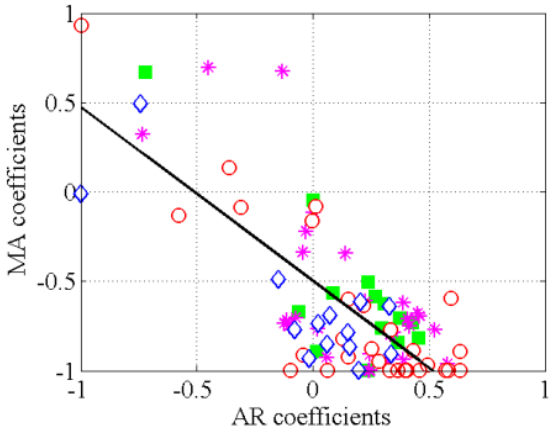


(d) ACF and PACF of residual

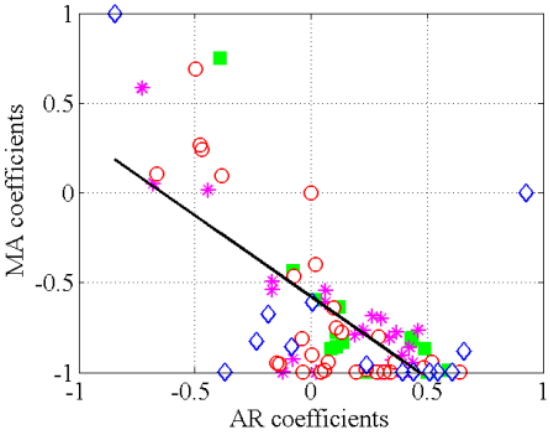
Figure 3.12 Application of ARIMA(1,1,1) model in Day 7

3.2.3 AR-MA coefficient plots from February to May, 2013

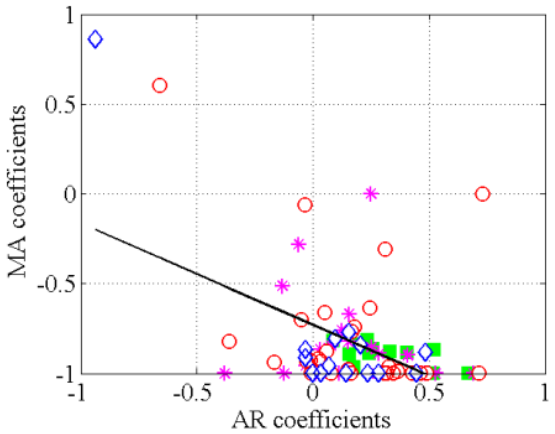
All estimated AR and MA coefficients from every one-day time-series data for all considering time-series GPS data were plotted. Figure 3.13(a-c) show the plots of the north tower point (#A), while the Figure 3.14(a) shows the center span point (#B) according the vertical direction (*z*-direction). In addition, the Figure 3.15 (a-b) are the plots of the quarter span point (#C) along to *x* and *z*-direction. Figure 3.16(a-c) are the plots of the south tower point (#D). In all figures, the plots are categorized by different markers depending on the each month data; where the green marker shows the coefficients in February, the pink marker shows them in March, the red circle marker and the blue diamond marker show them in April and May, respectively. Furthermore, the linear regression was also applied to plot in all figures, and the R-square values are shown below each plot.



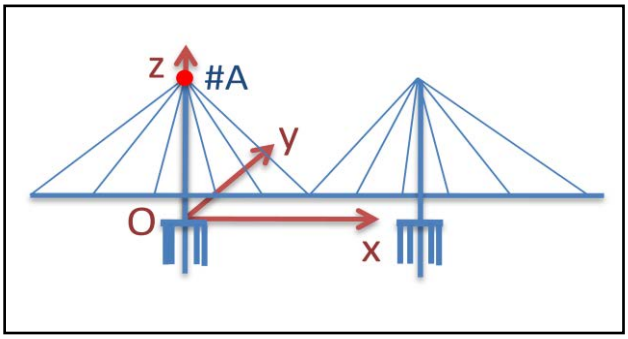
(a) *x*-direction at #A (R^2 = 0.578)



(b) *y*-direction at #A(R^2 = 0.479)

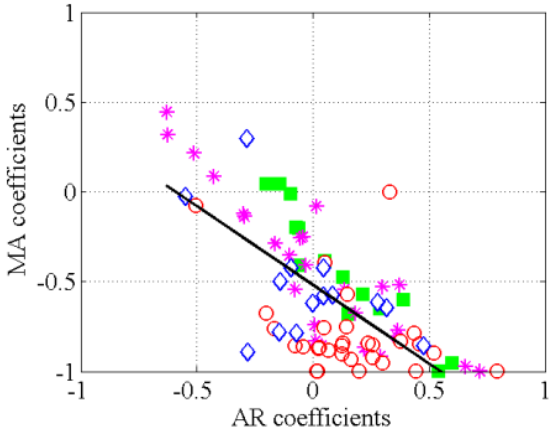


(c) *z*-direction at #A (R^2 = 0.221)

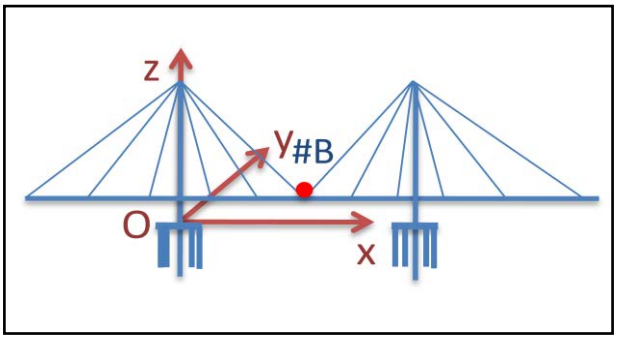


(d) Location of north tower point

Figure 3.13 AR-MA coefficient distributions of north tower

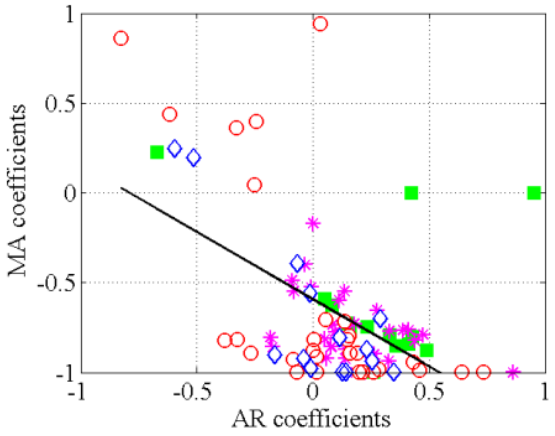


(a) *z*-direction at #B (R^2 = 0.486)

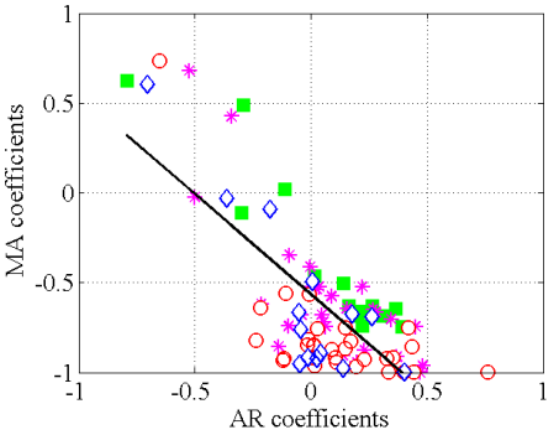


(b) Location of center span point

Figure 3.14 AR-MA coefficient distributions of center span

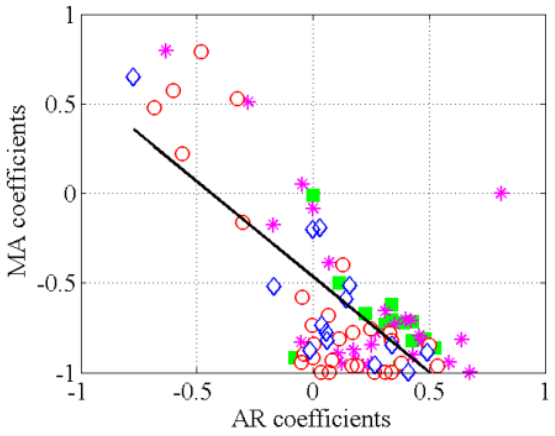


(a) x-direction at #C ($R^2 = 0.281$)

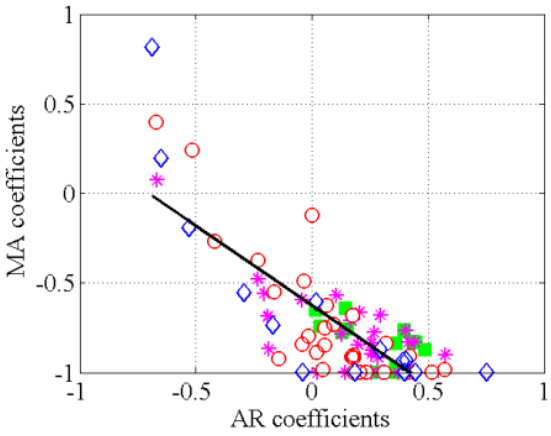


(b) z-direction at #C ($R^2 = 0.557$)

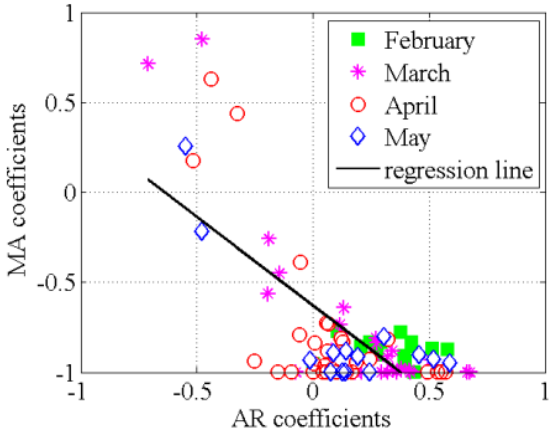
Figure 3.15 AR-MA coefficient distributions of quarter span



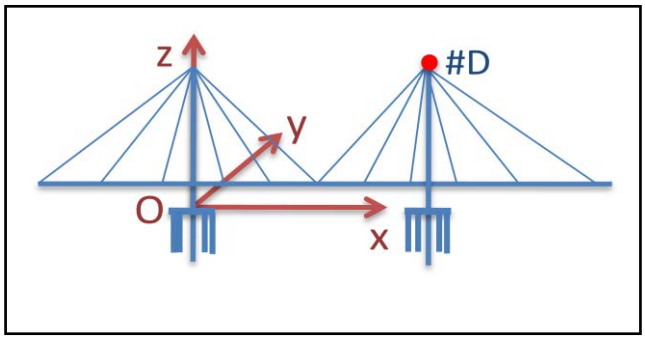
(a) x-direction at #D ($R^2 = 0.505$)



(b) y-direction at #D ($R^2 = 0.641$)



(c) z-direction at #D ($R^2 = 0.519$)



(d) Location of south tower point

Figure 3.16 AR-MA coefficient distributions of south tower

3.2.4 AR-MA coefficient plots in all 2013

To analysis further, the ARIMA(1,1,1) model was also applied for all GPS time series data in 2013. Four interest points with the considering directions were studied as mentioned above. For the case of less than four missing data occurrence, the least square approximation method was applied to interpolate the missing values in this study. Otherwise, if any day data set has more than three missing data, it was ignored in using ARIMA model. It means that the derived coefficients are not affected by the missing data. The AR-MA coefficients were derived from the ARIMA(1,1,1) model estimation to figure out as shown from Figure 3.17 to Figure 3.20, whereas the coefficients derived from longitudinal direction (*x*-direction), lateral direction (*y*-direction) and vertical direction (*z*-direction) were displayed by green square symbol, blue star symbol and red circle, respectively.

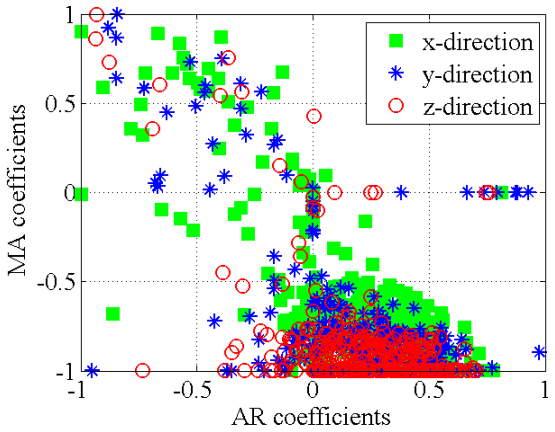


Figure 3.17 AR-MA coefficient distributions of north tower in all 2013

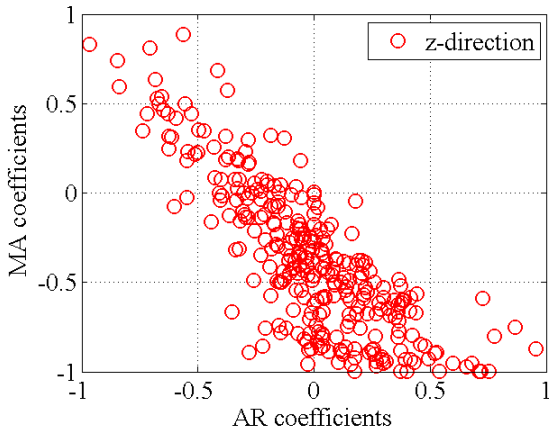


Figure 3.18 AR-MA coefficient distributions of center span in all 2013

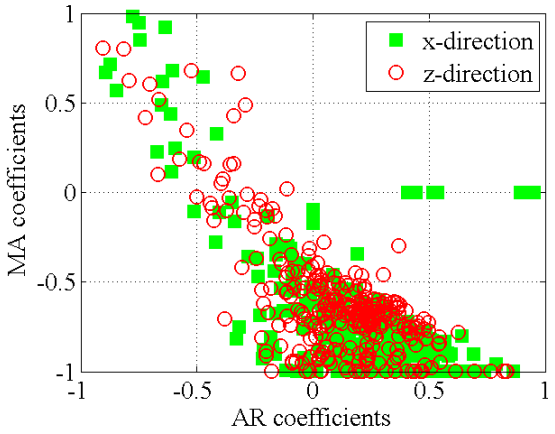


Figure 3.19 AR-MA coefficient distributions of quarter span in all 2013

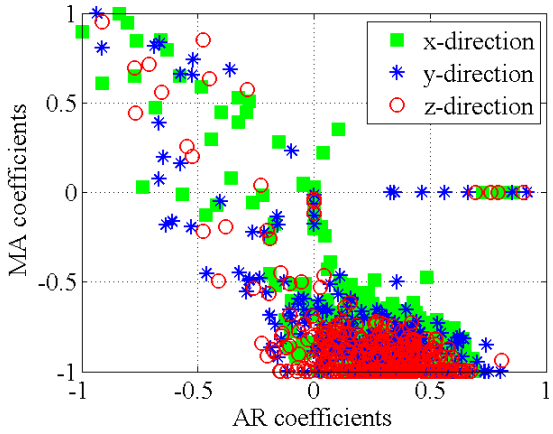


Figure 3.20 AR-MA coefficient distributions of south tower in all 2013

3.2.5 Discussion

Before the discussions here, it should be noticed that any significant structural condition changes had not been reported during this period in the *Can Tho* bridge. For the case of almost three month of 2013 (from February 15th to May 15th), relatively high *R*-square values are shown in the plots from the GPS data in the *z*-direction at both the center and quarter spans (#B and #C), those in the *x*-direction at the top of two towers (#A and #D), and those in all directions at the top of South tower (#D). In those data, the correlation coefficients with the temperature also showed high values, as mentioned in Table 2.4 (Chapter 2). Therefore, those regressions are considered to indicate the pattern of the global deformation due to the one-day periodic temperature change statistically.

In the case of all year 2013, the AR-MA coefficients of almost one year have a trend in their distribution. Figure 3.17 and Figure 3.20 show the coefficient plots of two tower points in all directions that almost coefficients distributed in the same field such as from -0.2 to 0.7 for AR coefficient and from -1 to -0.5 for MA coefficients. This kind of trend looks to the same in the Figure 3.19 which showed the coefficient plot of the quarter span point in longitudinal and vertical directions. On the other hand, Figure 3.18 shows the coefficient plot of the centre span point in vertical direction (*z*-direction) that showed clearly the distributed trend in orientation from north-west to south-east. Looking at the vertical direction distribution (red circle symbol) of the Figure 3.19, it has the same trend to compare with the Figure 3.18. It can be concluded that the girder points (#B and #C) have the same trend in movement according to vertical direction. This point also can be understood as one aspect of the global deformation of the target bridge. However, this kind of characteristic did not show in the vertical direction of the two tower points even though there are several coefficients scattered. It can be explained that the two tower points do not move so much in the vertical direction.

In both two cases of the coefficient plots, there are several coefficients that scattered in almost directions. They might were caused by noise and abnormal data. Therefore, any outlier detection procedure should be adopted for the statistical structural condition diagnosis; however, it can be said that those regressions in the AR-MA coefficients from the ARIMA model can be used as the base distributions there. It was then concluded that the global deformations due to the one-day periodic temperature change could be used for the global structural condition assessment.

3.3 Summary

The main content of this chapter is how to apply the Autoregressive Integrated Moving Average (ARIMA) time series model for the GPS observed data of the target bridge. First of all, the first part (section 3.1) presents the basic theory of time series analysis as well as how to apply this theory for each case of process is also mentioned. The second part shows the process of applying the ARIMA model into the target GPS data according to the interest points (section 3.2). The important results of this estimation clearly showed the global deformation of the target bridge that was mentioned in previous chapter, and the distribution of AR-MA coefficients would be used as a base for any outlier detection procedure to assess structural conditions.

Chapter 4

Numerical analysis of cable-stayed bridge model due to thermal changes

4.1. Introduction

Chapter 2 and chapter 3 are the analytical strategies of GPS application in structural health monitoring of the real cable-stayed bridge. The results showed that, the global deformation modes of the target bridge are clearly shown under the effects of the atmospheric temperature changes, and there is the same one-day periodic in the GPS time-series data and the atmospheric temperature data. Then, the Autoregressive Integrated Moving Average (ARIMA) model was applied for each one-day of GPS time-series data to extract the AR-MA coefficients. The distribution plots of those coefficients showed clearly the trend in their distribution. Even though there is no significant change of the target bridge that has been reported, there are some scattered coefficients in the plots. This kind of scattered coefficients may be caused by the effects of noise and abnormal GPS data.

Nowadays, modern bridge infrastructure often comprises of concrete and steel structures. Under operation of bridges, there are various factors that can affect to change the bridge structure conditions, such as the over loading due to increase in wheel loads of traffics, or effect of environmental source which cause corrosion of cables or mar the properties of boundary conditions. Construction stage with poor quality and lack of regular maintenance leads to major retrofit in bridge structures. Analysis of a finite model of cable stayed bridge under the effects of operational and environmental factors is an important issue that various researchers focused on. The dynamic behavior of a combined cable system of bridges under moving loads was investigated by a mathematical model that proposed in [35]. In addition, the stochastic perturbation technique and Monte Carlo Simulation method were used in analysis to present stochastic seismic finite element analysis of cable stayed bridge material properties by random fields [36]. The experiments on corroded galvanized steel wires at different corrosion levels showed that the wires were fractured by the mixed effects of corrosion and cyclic stresses [37]. Ronal (2000) [38] concluded that reduction in strength of the cable due to deterioration increases

with increase in dead load. Method of separating the wire into corrosion stages was given to calculate the strength of the cable that was used to analyze effect of cable degradation on dynamic behavior of cable stayed bridges [39].

In this chapter, a cable-stayed bridge model is built in the MIDAS civil software to validate its global deformation behaviors under the effects of one-day time-series temperature change as well as the ability of using AR-MA coefficients to detect the structural condition changes. Firstly, the one-day time-series temperature data is used as the input data for model to acquire the time-series displacements of whole model in two cases of study: normal condition and structural condition change. Then, the acquired time-series displacements of the interested points considering the global deformation modes are applied the ARIMA model to extract the AR-MA coefficients. The AR-MA coefficients in both cases: normal condition and structural condition change are plotted to show their distribution in comparison.

4.2. Analytical model description

4.2.1. Modelling of cable-stayed bridge

Three dimensional type of cable-stayed bridge is referred to model in normal condition by using the finite element software Midas Civil [39]. The total length of the cable-stayed bridge model is 420m. It has three continuous spans of cable-stayed bridge of 100m each at the ends and 220m in the center. The width of the bridge is 15.6m of two lanes. Two towers with overall height of 90m support the structure by means of two planes of cables in fan type arrangement. Spacing of the cable is 10m in the main span. The general layout of the cable-stayed bridge model is shown in Figure 4.1. The material properties and sectional properties of structural elements were referred from a tutorial on Midas Civil software as figured out in Table 4.1 and Table 4.2.

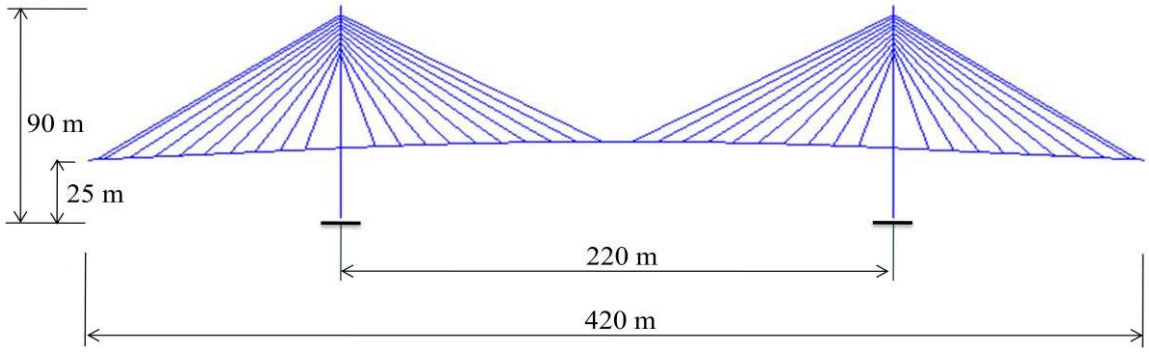


Figure 4.1 General layout of the model

Table 4.1 Material properties of the model

Material ID	Name	Modulus of Elasticity (kN/m^2)	Poisson's ratio	Weight Density (kN/m^3)
1	Cable	1.9613×10^8	0.3	77.09
2	Girder	1.9995×10^8	0.3	77.09
3	Pylon	2.78×10^7	0.2	23.56
4	CBeam-Girder	1.9995×10^8	0.3	77.09
5	CBeam-Pylon	2.78×10^7	0.2	23.56

Table 4.2 Section properties of the model

Section ID	Name	Area (m^2)	I_{xx} (m^4)	I_{yy} (m^4)	I_{zz} (m^4)
1	Cable	0.0052	0.0	0.0	0.0
2	Girder	0.3092	0.007	0.1577	4.7620
3	Pylon	9.2	19.51	25.567	8.123
4	CBeam-Girder	0.0499	0.0031	0.0447	0.1331
5	CBeam-Pylon	7.2	15.79	14.472	7.992

In modeling, the height of the pylon above the deck is about 60m, and the location of pylon is 100m from each end. The arrangement of cables in the longitudinal direction in fan type and the transverse arrangement is of inclined plane system. The length of the cables varies from 43.04m (the 10^{th} cable – nearest to the tower) to 114.66m (the 1^{st} cable – farthest from the tower).

Therefore, 40 cables are distributed on four sides of both the towers symmetrically. The pre-calculated initial tension was built in configuration in software which was used to account for the geometric stiffness. The 3D model which is built in software, and the interest points considering the global deformation are presented in Figure 4.2.

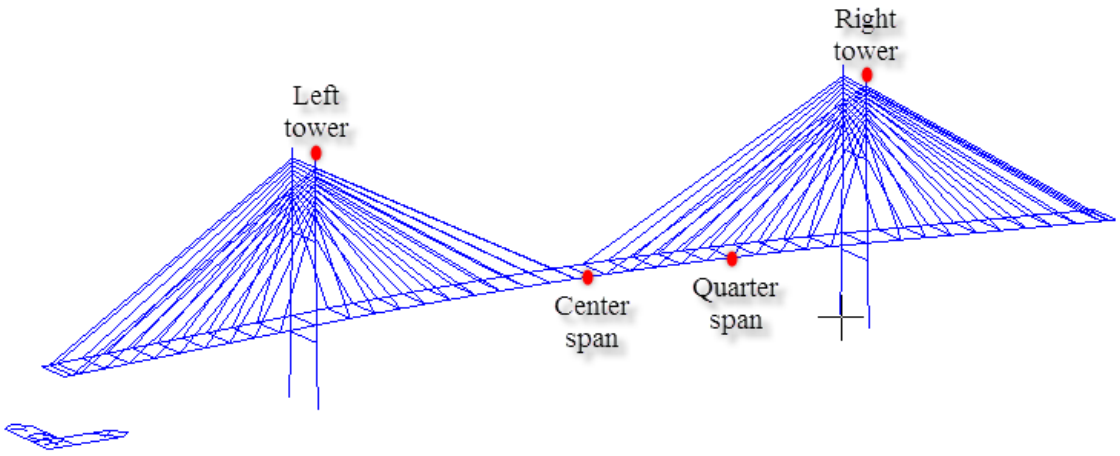


Figure 4.2 Numerical model and interest points

To analyze the model on software, the boundary conditions are very important. In this case of numerical, the bearings are generated at two towers and two end piers of the model as shown an example in Figure 4.3 , where the bearing properties are follows: SDx = 199,736,032 kN/m; SDy = 73,373 kN/m; and SDz = 73,373 kN/m, respectively. In addition, the boundary supports for the analytical model are as follows (Figure 4.4 and Figure 4.5):

- Tower base, Pier base: Fixed condition at (Dx, Dy, Dz, Rx, Ry, Rz)

- Connections between Main Girders and Bearings: Rigid links are fixed at (Dx, Dy, Dz, Rx, Ry, Rz)

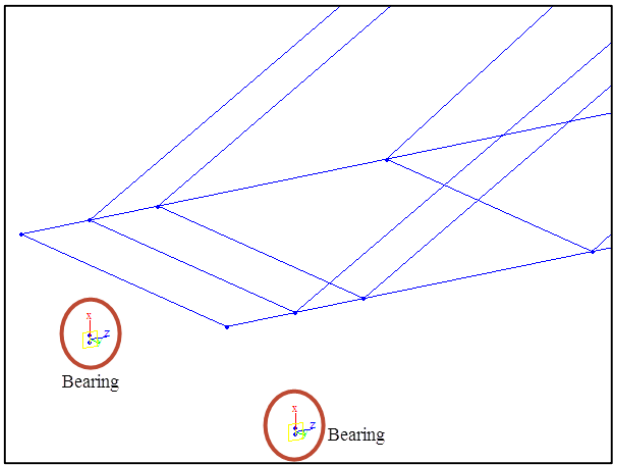


Figure 4.3 Bearing generation

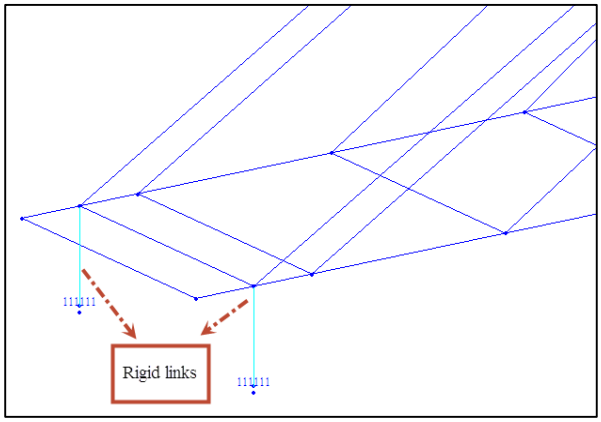


Figure 4.4 Connections between main girders and pier bearing using rigid links

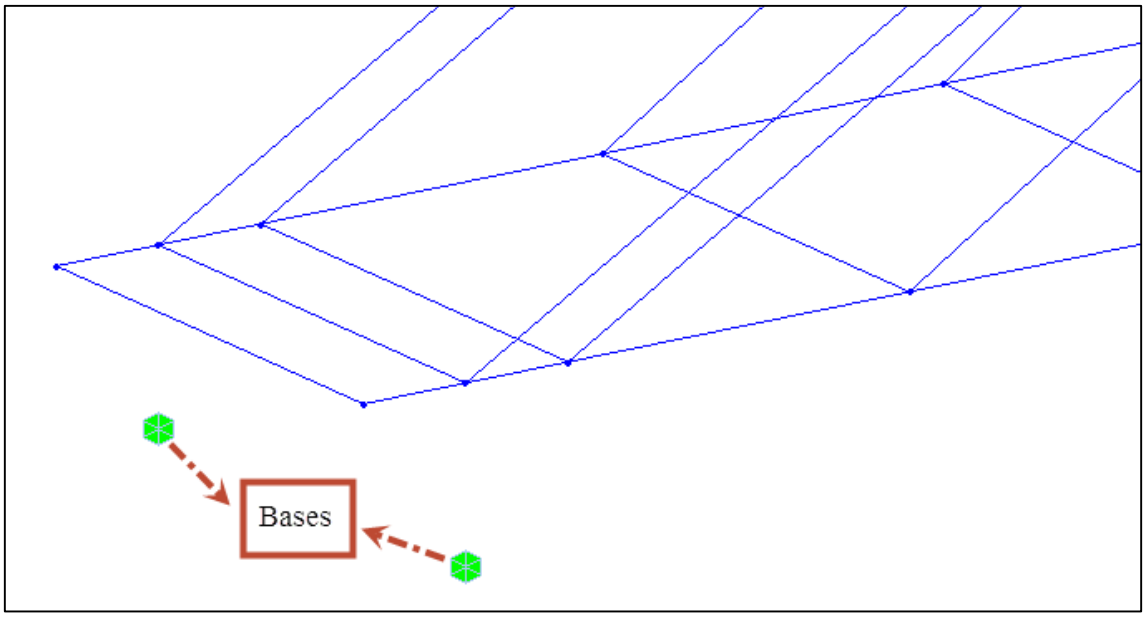


Figure 4.5 Base generation

4.2.2. The input temperature data and time-series displacement acquisition

Here, the ten-minute average temperature data of one day in sinusoidal type were generated with the amplitude of 2 Celsius as shown in Eq. (4.1) and figured out in Figure 4.6. The coordinates of all nodes and all cable tensions of the model are updated step by step of each case of temperature changes based on the nonlinearity analysis. The displacements of all nodes in both cases: normal condition and structural condition change are then acquired due to the effect of

temperature changes. In the case of the structural condition change, the typical bearing properties of boundary conditions as well as the tension of some cables are selected to change the value. In detail, the bearing properties at two towers and one of anchorage are changed to nearly fix conditions by increasing the value of stiffness direction along to z-direction and the value of stiffness rotation along to y-direction. In addition, the three selected cable tensions are reduced of 5%, 10% or 25% and so on. The case studies of the structural condition change are figured out in the Figure 4.7, and summarized in the Table 4.3 and Table 4.4.

$$temp = A\sin\left(\frac{2\pi}{T}t\right) + 27 \tag{4.1}$$

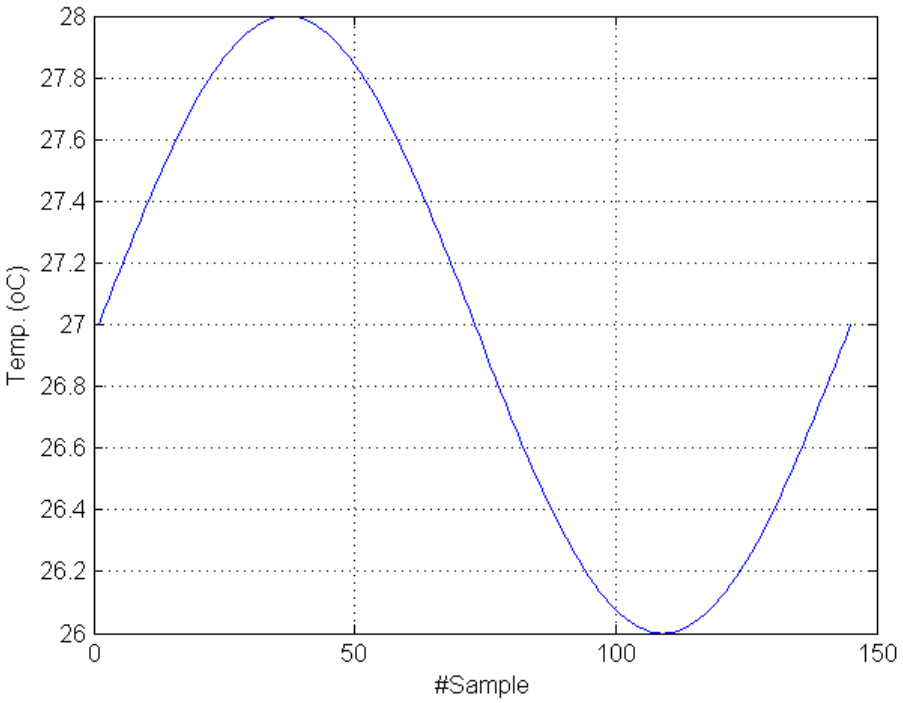


Figure 4.6 Time series temperature data of 1 day

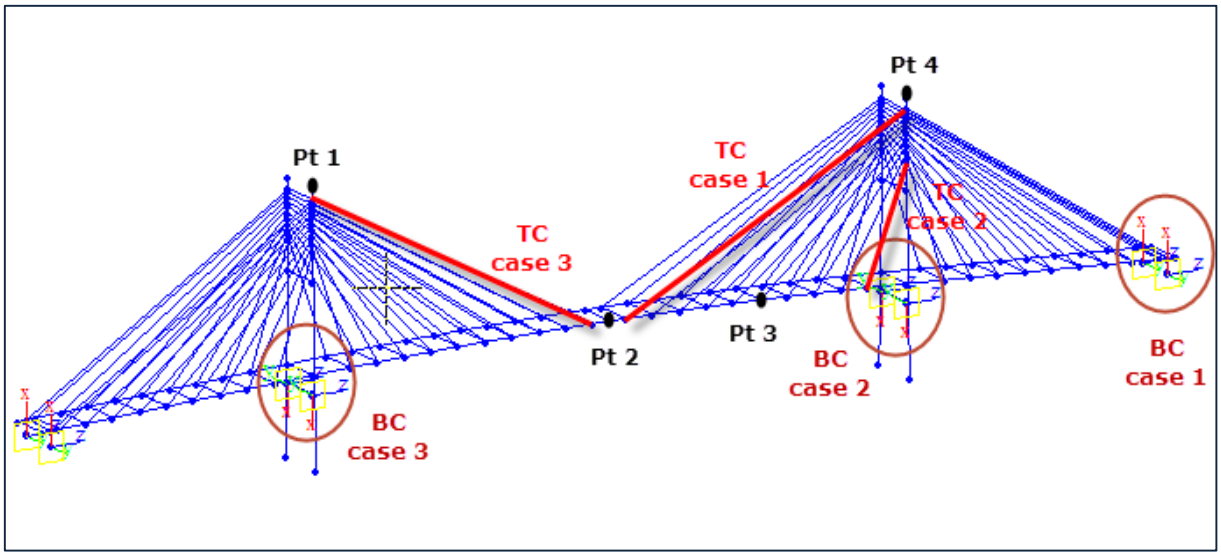


Figure 4.7 Layout of structural condition change

In the case of changing the boundary conditions, the bearing properties of the right tower (see locations in figure 4.7) was analyzed to show the different effects in two cases of increasing the stiffness direction and the stiffness rotation (BC case 2 and BC case 4 in Table 4.3). This kind of test was also studied in the case of all towers (BC case 5 and BC case 5-1 in Table 4.3). Additionally, there are two highest tension cables and one lowest tension cable that are choose to analyze the effect of tension reductions into the structural condition change which can be predicted by the distribution of AR-MA coefficients later. Especially, the tension change in case 1 was investigated in four different levels of reduction (see Table 4.4).

Table 4.3 Boundary condition change

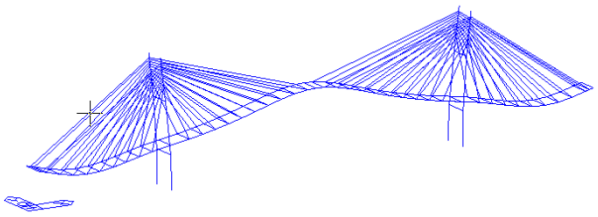
Case study	Bearing property change	Notice
BC case 1	SDz: $*10^4$ SRy: $*10^5$	At the anchorage
BC case 2	SDz: $*10^4$ SRy: $*10^5$	At the right tower
BC case 3	SDz: $*10^4$ SRy: $*10^5$	At the left tower
BC case 4	SDz: $*10^{20}$ SRy: $*10^{20}$	At the right tower
BC case 5	SDz: $*10^6$ SRy: $*10^8$	At the all towers
BC case 5-1	SDz: $*10^{20}$	At the all towers

	SRy: $*10^{20}$	

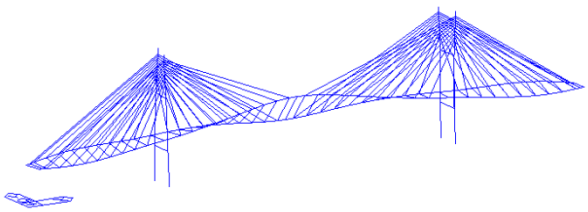
Table 4.4 Cable tension reduction

Case study	Portion of reduction	Notice
TC case 1	5% 10% 25% 50%	Tension change in case 1
TC case 2	10%	Tension change in case 2
TC case 3	10%	Tension change in case 3

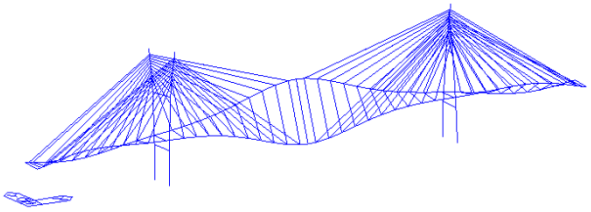
Additionally, the first 6th mode shapes of numerical finite element models are shown in some cases of analysis, such as normal condition (Figure 4.8), boundary condition changes in case 2 and case 4 (Figure 4.9 and Figure 4.10) (same location, but difference in increasing stiffness direction and stiffness rotation of bearing properties), boundary condition changes in case 5-1 (Figure 4.11) (increase high values of bearing properties in all towers), and cable tension reduction of 10% in case 1 (Figure 4.12). The frequency components of the first 6th mode shapes are also summarized in the Table 4.5.



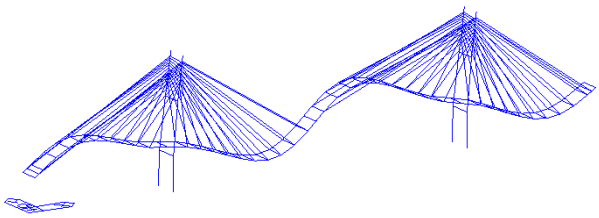
(a) 1st mode



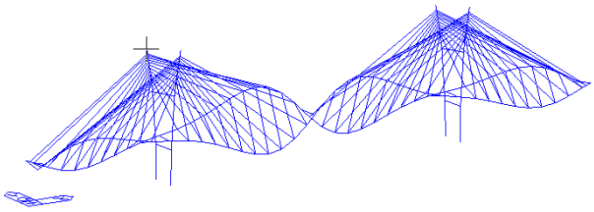
(b) 2st mode



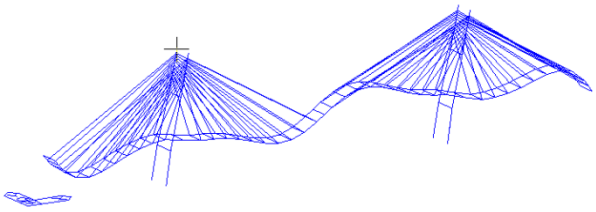
(c) 3st mode



(d) 4st mode

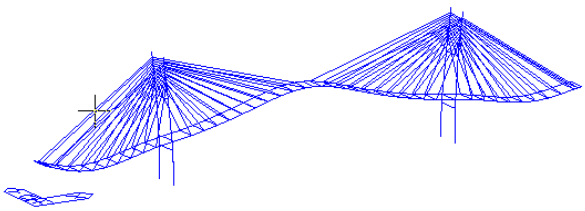


(e) 5st mode

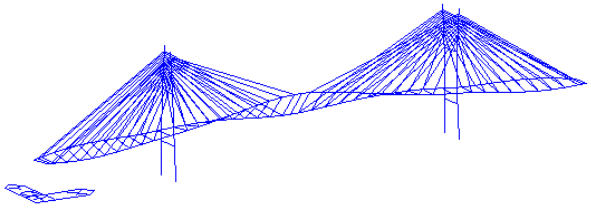


(f) 6st mode

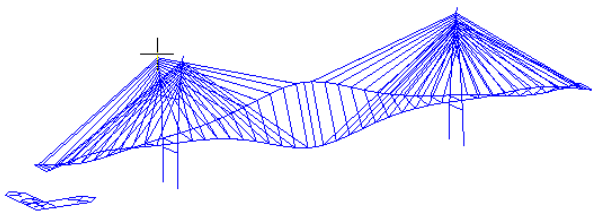
Figure 4.8 Mode shapes of model in normal condition



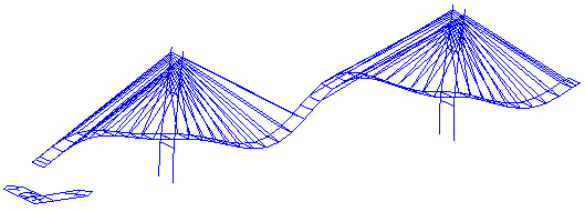
(a) 1st mode



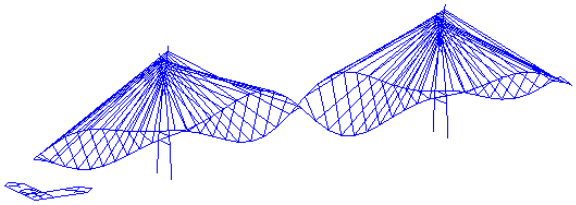
(b) 2st mode



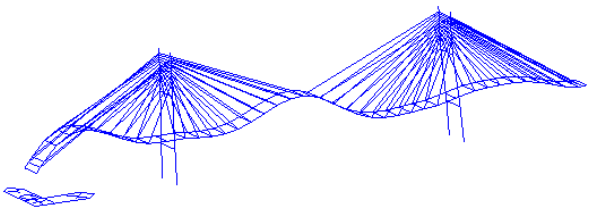
(c) 3st mode



(d) 4st mode

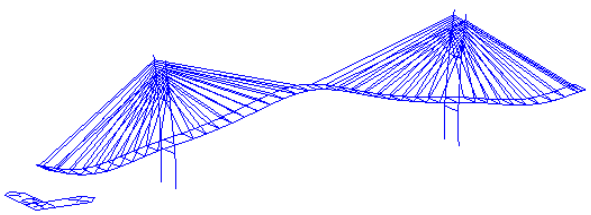


(e) 5st mode

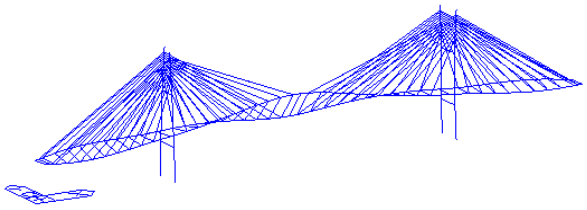


(f) 6st mode

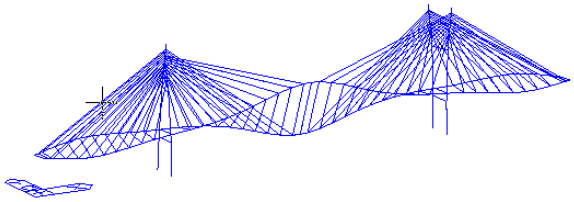
Figure 4.9 Mode shapes of model in BC case 2



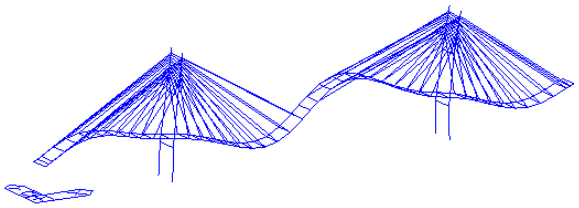
(a) 1st mode



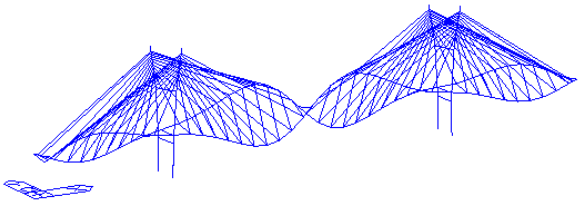
(b) 2st mode



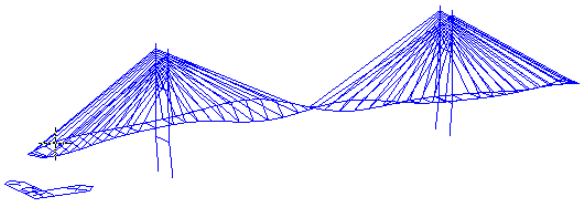
(c) 3st mode



(d) 4st mode

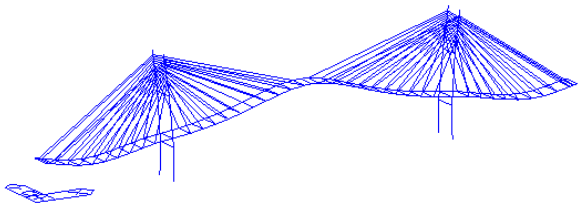


(e) 5st mode

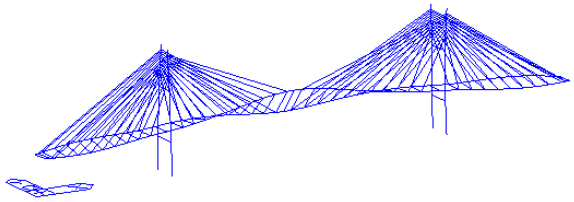


(f) 6st mode

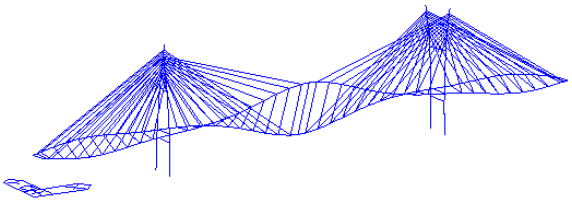
Figure 4.10 Mode shapes of model in BC case 4



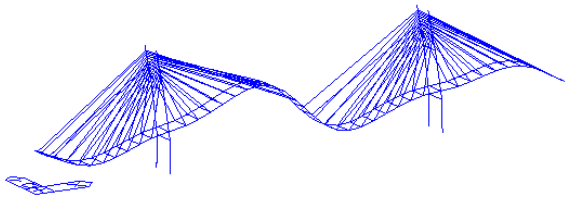
(a) 1st mode



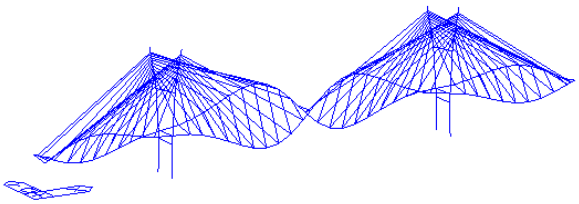
(b) 2st mode



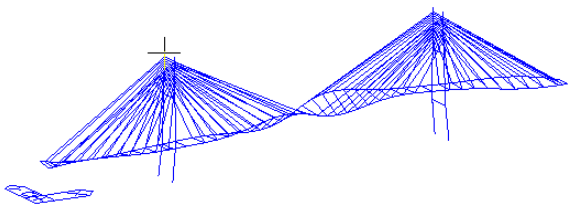
(c) 3st mode



(d) 4st mode

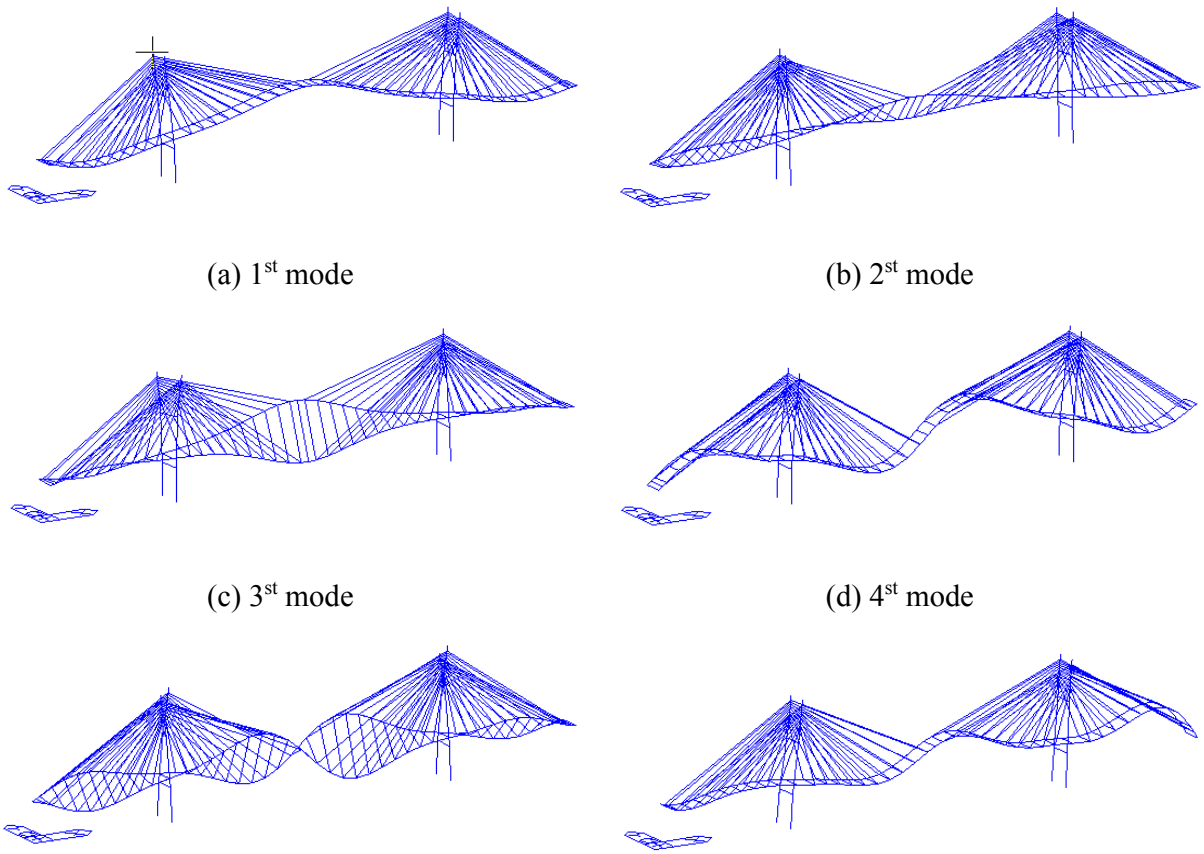


(e) 5st mode



(f) 6st mode

Figure 4.11 Mode shapes of model in BC case 5-1



(a) 1st mode (b) 2st mode

(c) 3st mode (d) 4st mode

(e) 5st mode (f) 6st mode

Figure 4.12 Mode shapes of model in TC case 1-10%

Table 4.5 Frequency components of the first six modes

#Mode	Normal condition	BC case 2	BC case 4	BC case 5-1	TC case 1 -10%
1	1.174	1.199	1.208	1.269	1.191
2	1.199	1.211	1.256	1.315	1.199
3	1.496	1.503	1.507	1.517	1.514
4	1.652	1.731	1.739	1.781	1.654
5	2.020	2.043	2.051	2.080	2.053
6	2.318	2.484	2.488	2.489	2.455

It can be seen that, when the structural conditions change, i.e. boundary condition changes, or cable tension reductions, the frequency components of mode shapes of model increase slightly.

Looking at the mode shape figures, the 4th mode shapes seem to not change in shape of model, however, the last two mode shapes look more change to compare with the mode shapes in normal condition.

In the numerical analysis, four interested points are two tower points, center span point, and quarter span point as shown in Figure 4.7. The selection of these interested points is based on the global deformation behavior that was analyzed in the previous chapters of real cable-stayed bridge. Figure 4.13 shows the strategy of time series displacement acquisition of the numerical model, whereas the ten-minute average temperature data is used as input into the numerical model to acquire the time-series displacements of four interested points. The output time-series displacement data in one day of four interested points under the effects of input temperature change were taken as example to show in the different cases from Figure 4.14 to Figure 4.17.

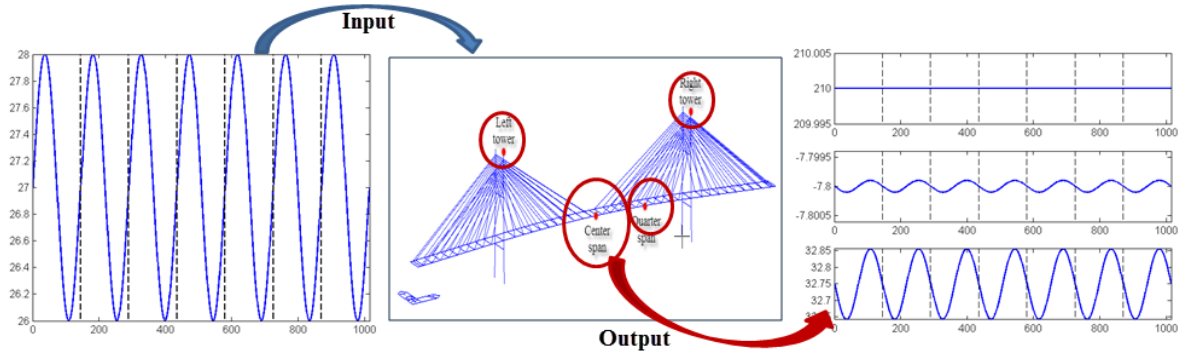
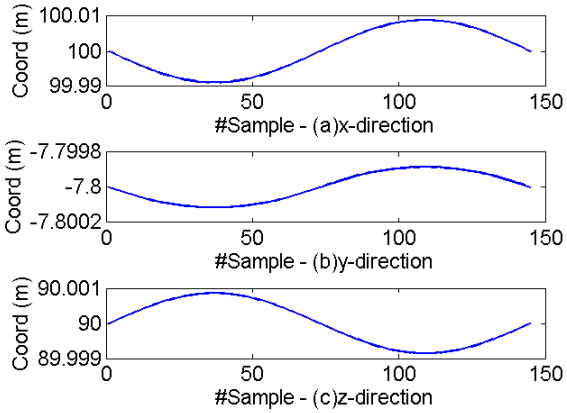
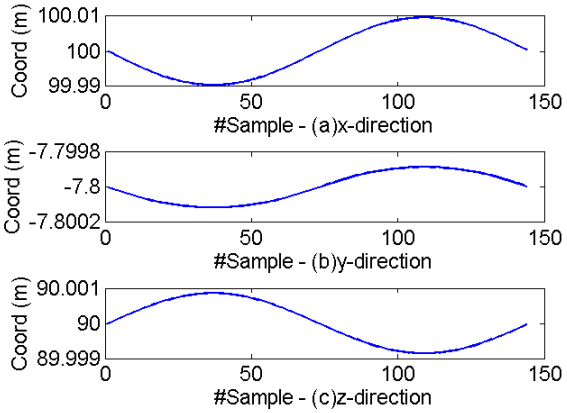


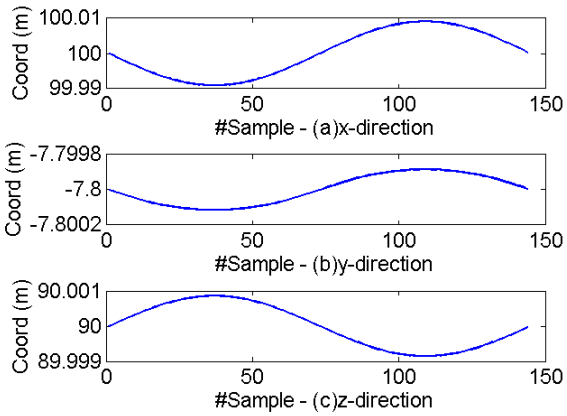
Figure 4.13 Strategy of time-series displacement acquisition of numerical model



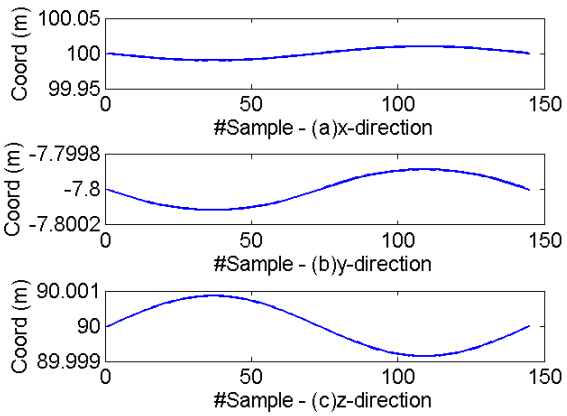
(a) Normal condition



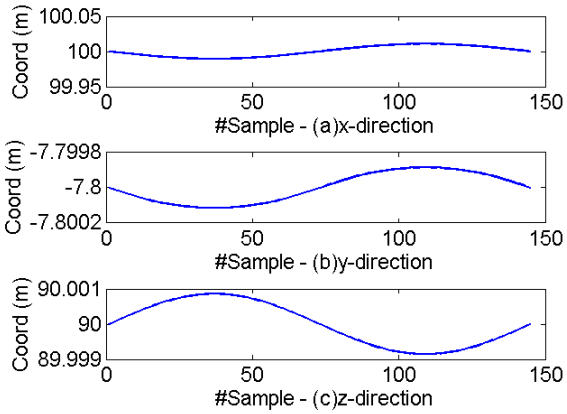
(b) Tension change 10% in case 1



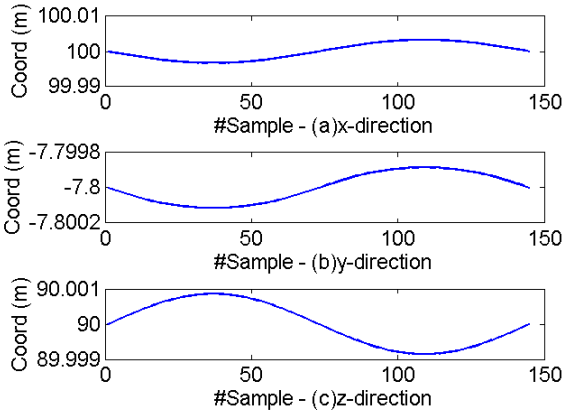
(c) Boundary condition change in case 2



(d) Boundary condition change in case 4

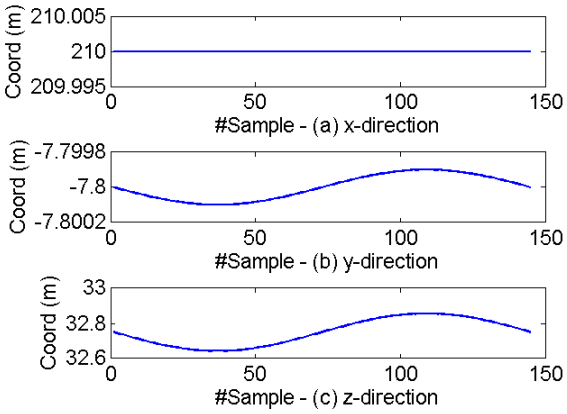


(e) Boundary condition change in case 5

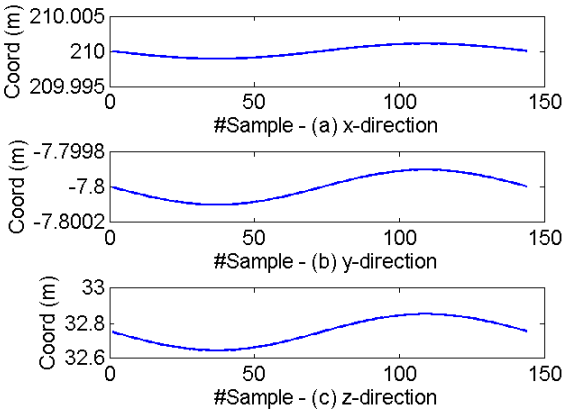


(f) Boundary condition change in case 5-1

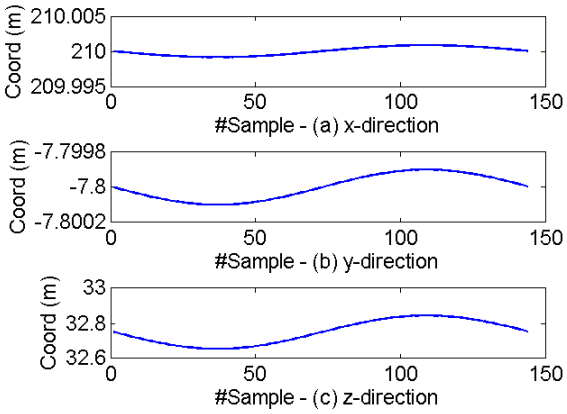
Figure 4.14 The output time-series displacement data of left tower point



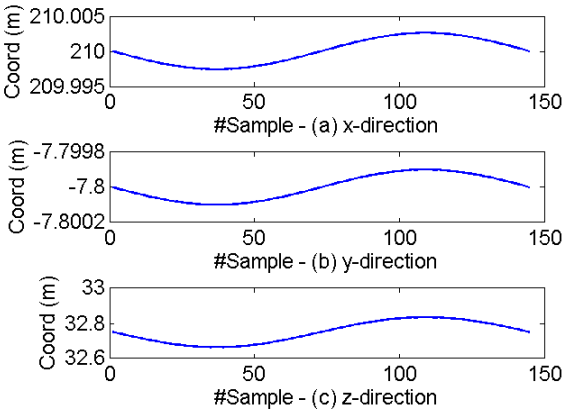
(a) Normal condition



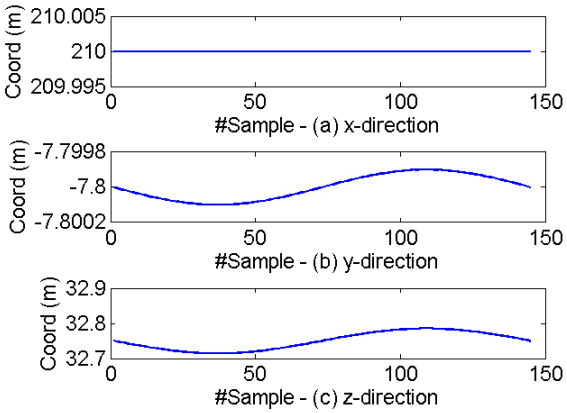
(b) Tension change 10% in case 1



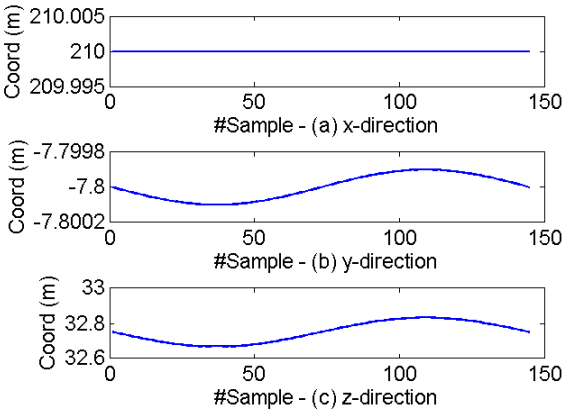
(c) Boundary condition change in case 2



(d) Boundary condition change in case 4

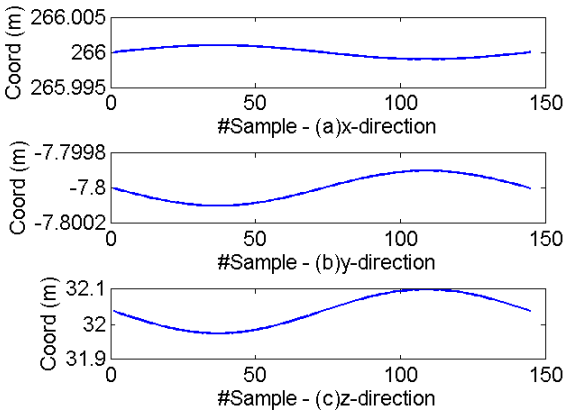


(e) Boundary condition change in case 5

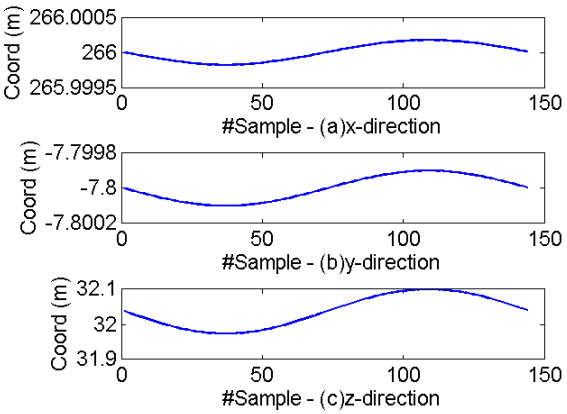


(f) Boundary condition change in case 5-1

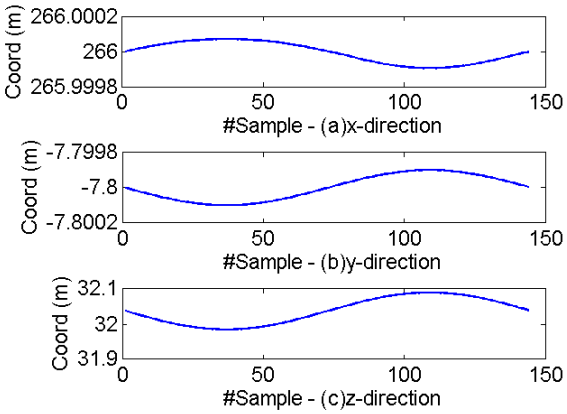
Figure 4.15 The output time-series displacement data of center span point



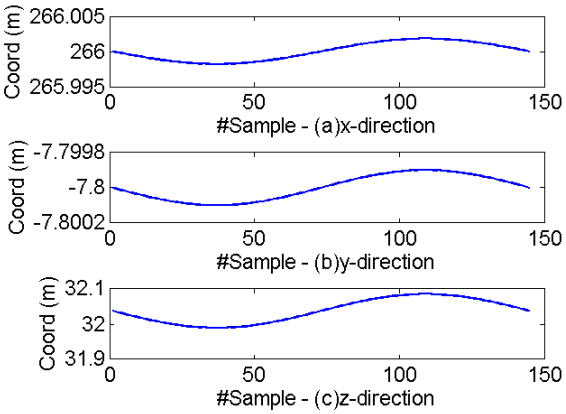
(a) Normal condition



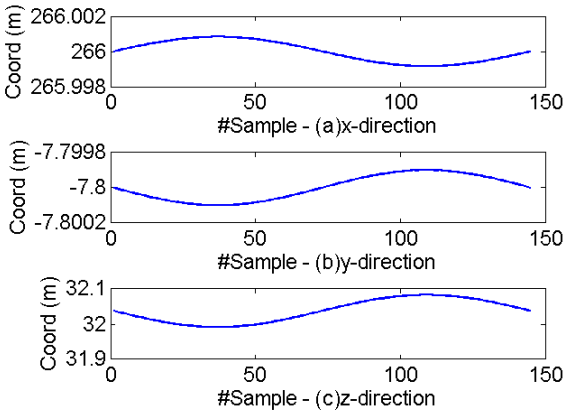
(b) Tension change 10% in case 1



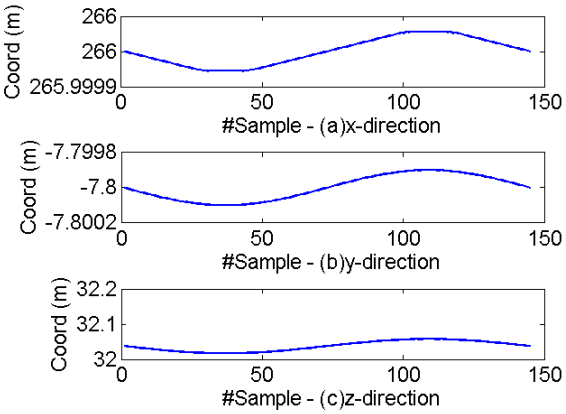
(c) Boundary condition change in case 2



(d) Boundary condition change in case 4

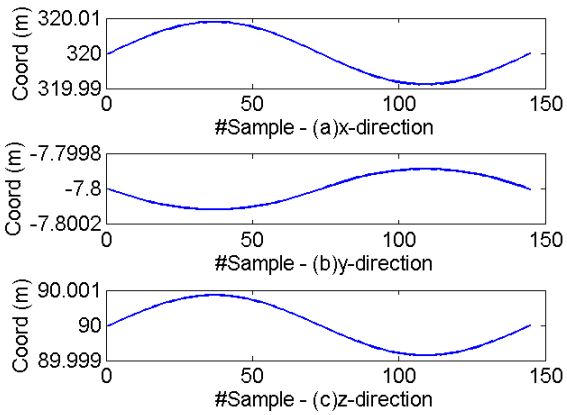


(e) Boundary condition change in case 5

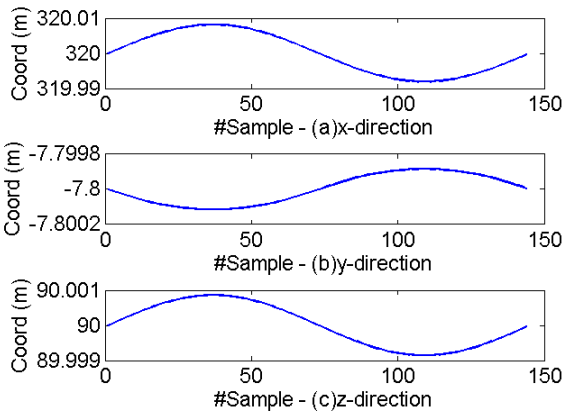


(f) Boundary condition change in case 5-1

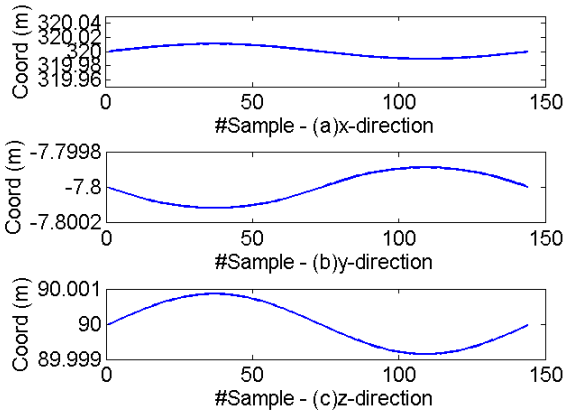
Figure 4.16 The output time-series displacement data of quarter center span point



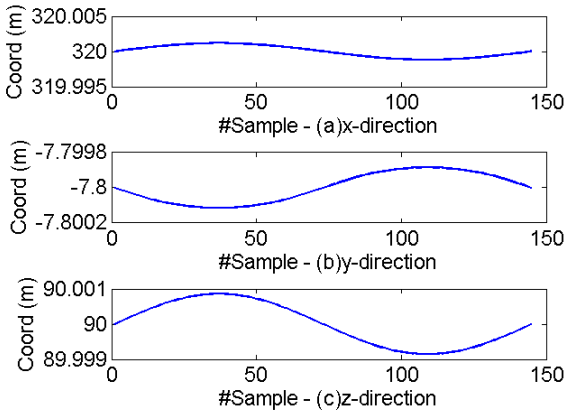
(a) Normal condition



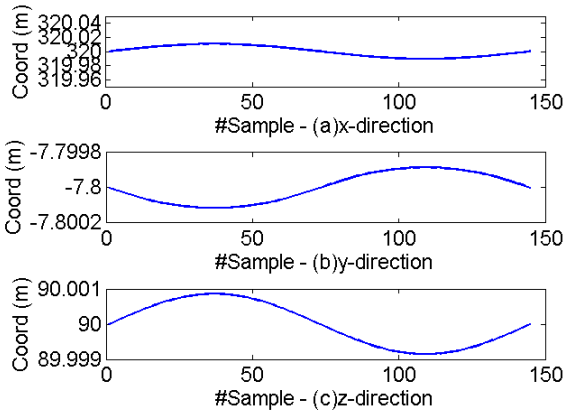
(b) Tension change 10% in case 1



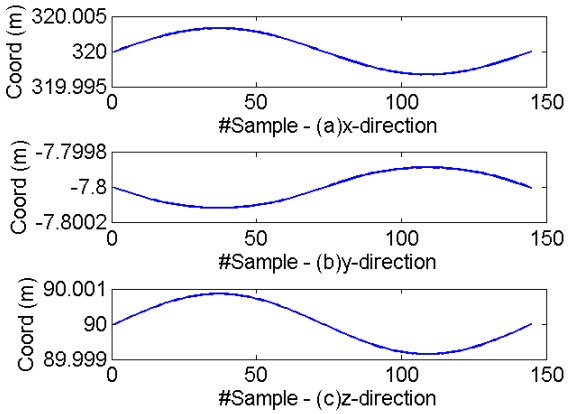
(c) Boundary condition change in case 2



(d) Boundary condition change in case 4



(e) Boundary condition change in case 5



(f) Boundary condition change in case 5-1

Figure 4.17 The output time-series displacement data of right tower point

There are some remarks which could be realized from the all plots. First of all, there is completely a periodic in the output time-series displacements due to the periodic of the input temperature data. In the cases of tower points, the displacement range along to *x*-direction (longitudinal direction) in all cases of those points that is much larger than other directions (Figure 4.14 & Figure 4.17). Therefore, the longitudinal direction is the significant direction of the numerical model. In addition, the displacement range along to *z*-direction (vertical direction) in all cases of the girder points is much larger than other directions (*x*- and *y*-directions), and the vertical direction of girder points is also the significant direction (Figure 4.15 & Figure 4.16). Looking at the plots of *x*-direction (longitudinal direction) of center span point in the case of normal condition and the boundary condition change at all towers (see Figure 4.15 (a, e, f), there is no deflection. It can be explained that the analytical model is completely symmetric model, thus

the center point does not move along to *x*-direction. However, the deflection along to *x*-direction is larger than *y*-direction (lateral direction) in the other cases such as the tension change 10% in case 1 (Fig. 4.15 (b)), or the boundary condition change in case 2 and case 4 (Fig. 4.15 (c-d)). It can be explained that the boundary condition changes and cable tension change in those cases were at non-symmetric locations. Furthermore, looking at the Figure 4.15 (c) and Figure 4.15 (d), the displacement range along to *x*-direction increases when the bearing properties of the right tower increase, while the displacement range along to *y*- and *z*-direction are almost same in two cases. In the case of boundary condition changes in all towers, when the bearing properties increase, the displacement range along to z-direction also increase, while the displacement ranges of other directions are almost same in two cases, as shown in the Figure 4.15 (e – f). The discussions in the center span point are also the same trends in other interested points. It can be concluded that, the displacements of four interested points along to the significant directions are changed when the structural conditions of model change.

4.2.3. Adding random white noise

The output acquired time-series data are then added the random white noise. The range of random noise is referred from the real cable-stayed bridge case where the ratio of the error and the displacement ranges in horizontal plan (*x*- and *y*-direction) is around 25% and this ratio along to vertical direction is around 12%. Therefore, the ranges of adding random noise are based on those ratios. Consequently, the Figure 4.18 shows the sample of the adding random noise, and the one week sample of numerical time-series data with random noise in the normal condition, the boundary condition change in case 2 and case 4 are shown in the Figure 4.19, Figure 4.20, and Figure 4.21, respectively.

Looking at the one week example of time-series data, the displacement range of *x*- and *y*-direction are very small to compare with *z*-direction. However, the displacement range along to *x*-direction is much larger than *y*-direction except those in the center span point. It can be seen that, the displacement range along to x-direction of all four points change when the boundary condition change. When the bearing properties increase (BC case 2 and case 4), the displacement ranges along to *x*-direction of the tower points decrease, while those of girder points increase. Considering the global deformation behavior of bridge model as referred the results of the real cable-stayed bridge (chapter 2), the significant directions of the bridge model are *x*-direction of tower points, *z*-direction of girder points, and the *x*-direction of the quarter span point, respectively. These significant directions are used to analyze in the next step of study.

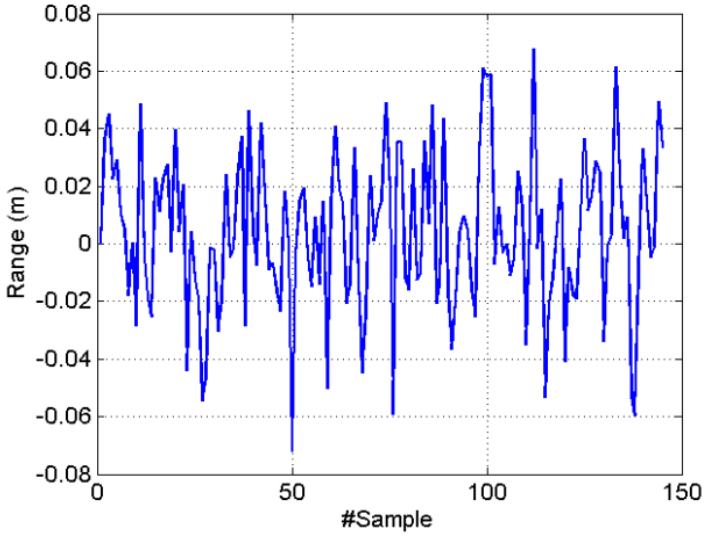
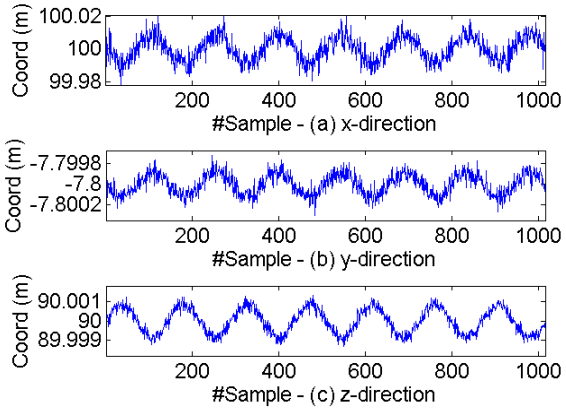
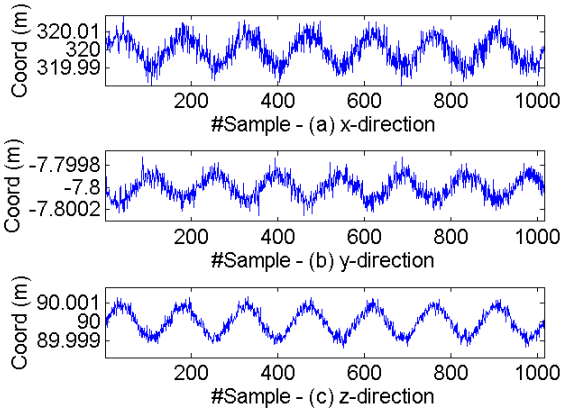


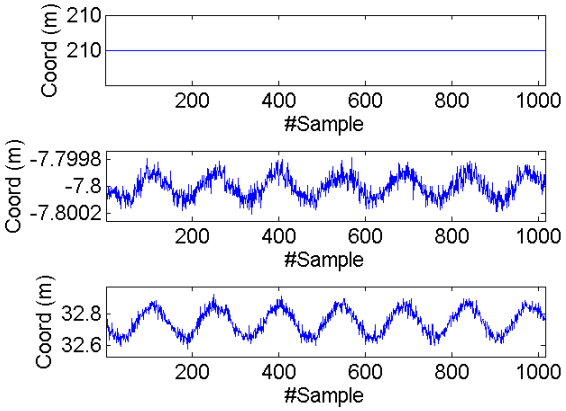
Figure 4.18 The random noise sample



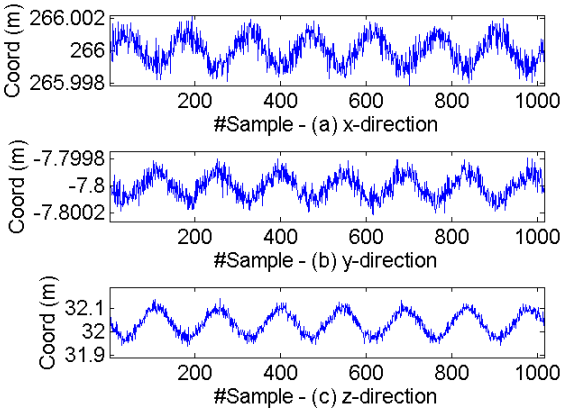
(a) The left tower point



(b) The right tower point

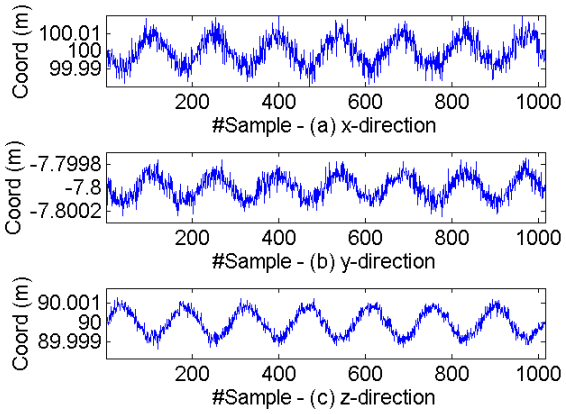


(c) The center span point

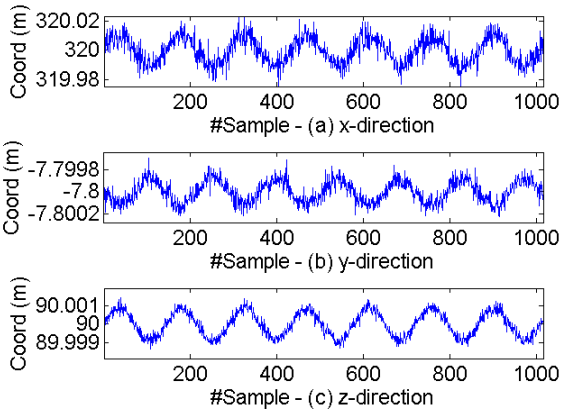


(d) The quarter span point

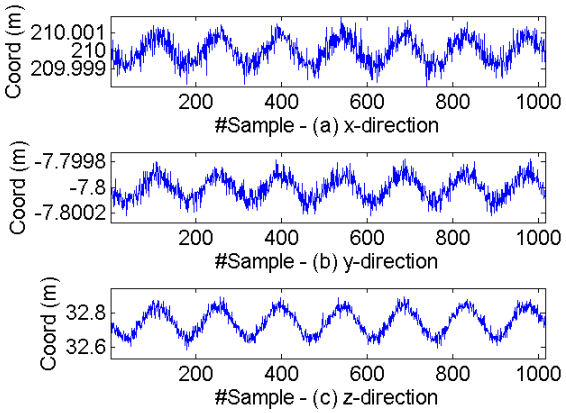
Figure 4.19 The one week data of the interested points in normal condition



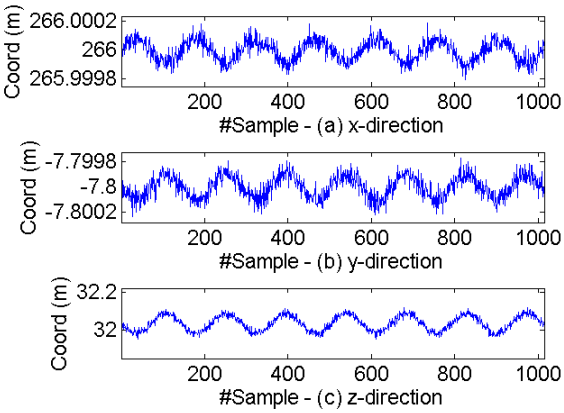
(a) The left tower point



(b) The right tower point

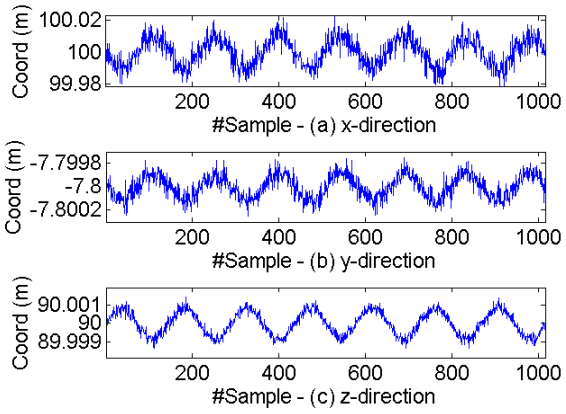


(c) The center span point

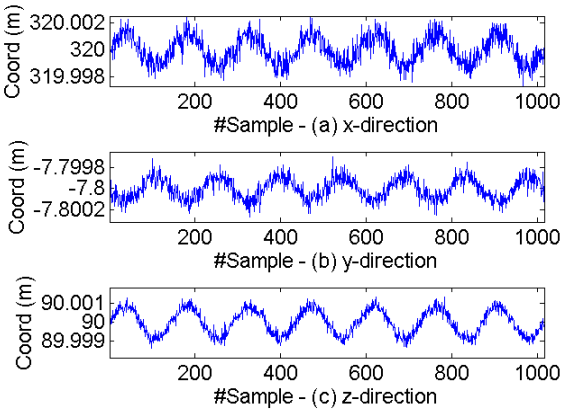


(d) The quarter span point

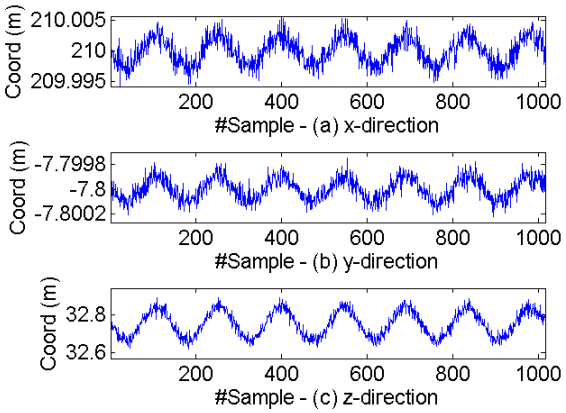
Figure 4.20 The one week data of the interested points in boundary condition change case 2



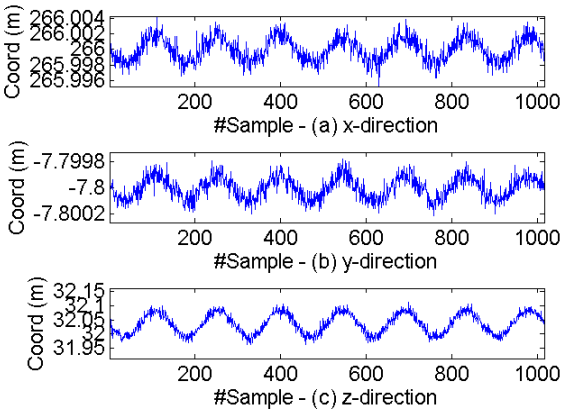
(a) The left tower point



(b) The right tower point



(c) The center span point



(d) The quarter span point

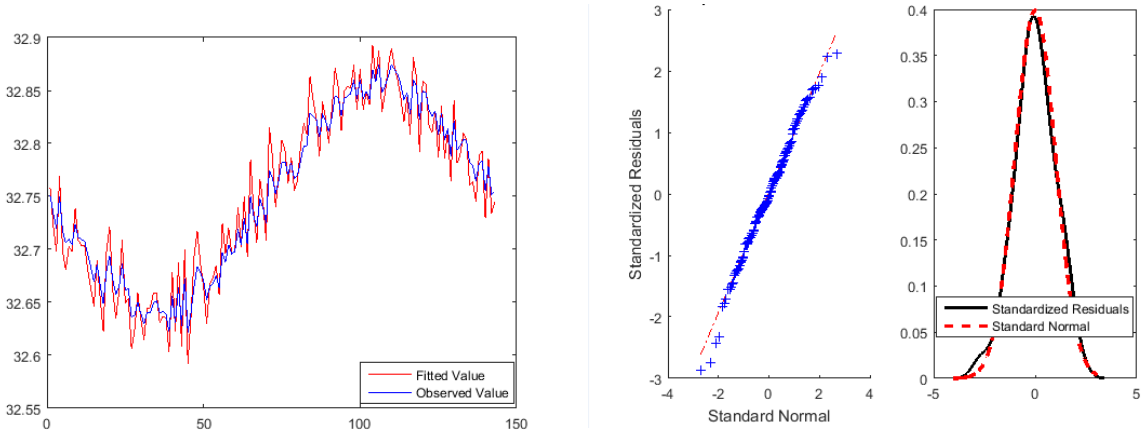
Figure 4.21 The one week data of the interested points in boundary condition change case 4

4.3. Application of ARIMA model for the numerical time-series data

The numerical time-series data in the normal condition case and the case of structural condition change were then applied the ARIMA model to extract the AR-MA coefficients. For the validation of the ARIMA model application in the GPS time-series, the ARIMA(1,1,1) was applied for the numerical analysis. Here, the length of time-series numerical data is taken in 20 days, and the data is divided to day by day for applying the ARIMA(1,1,1) model. Considering the global deformation behavior of numerical model, the *x*-direction (longitudinal) of the left and right tower points (#Pt1 and #Pt4), quarter span point (#Pt3), and the *z*-direction (vertical) of the center and quarter center points (#Pt2 and #Pt3) were taken to analyze in this step of study.

Figure 4.22 shows the example results of one-day ARIMA model performance that the time-series from the estimated ARIMA(1,1,1) model and the original one are overlaid in Figure 4.22(a), and the Figure 4.22(b) is the standardized distribution of the residual errors with the standardized normal distribution. It can be seen that, the estimated time-series data and the original one are fitted quite good, and the distribution of the residual error looks like the white noise distribution. To ensure the normal distribution of the residual error, the Ljung-Box Q-test (LBQ-test) [24] was also conducted to check whether the residual error distribution showed the white noise process with the normal distribution. The null hypothesis here was; "the residual was the white noise."In the results, the statistical possibility *p*-values at lags (3, 6, 9) were (0.1503, 0.2767, 0.0808) that are all larger than the 5% significant level, and the test statistic values are

calculated at (2.0695, 5.1049, 12.6617), which are all smaller than the critical values at 5% significant level (3.8415, 9.4877, 14.0671). It means that the null hypothesis was not rejected. The residuals are thus identified as the white noise process. We also checked those performances in the estimations to other one-day time-series and confirmed that the estimations were conducted with almost the same and appropriate accuracies.



(a) Overlay of estimated and numerical series (b) Distribution of the residual error

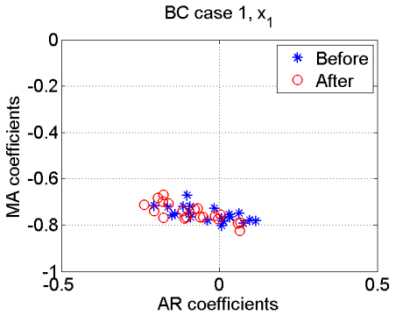
Figure 4.22 Estimated results of one ARIMA (1,1,1) model

4.4. Structural condition assessment of analytical model due to the AR-MA coefficient distributions

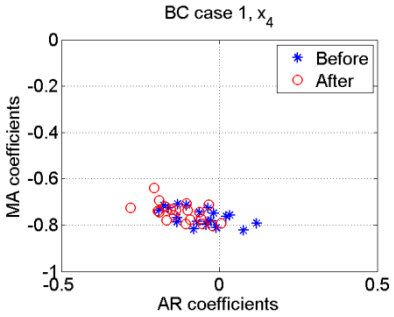
4.4.1. The case of boundary condition changes

4.4.1.1 Plots of AR-MA coefficients

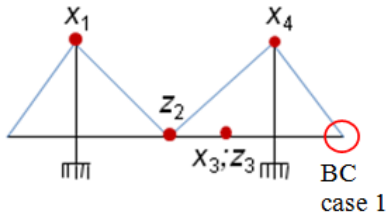
In this chapter, the AR-MA coefficients derived from the application of ARIMA model for the numerical data were figured out in each case of the boundary condition changes as shown from Figure 4.23 to Figure 4.28, continuously. In all plots, the blue star symbol denotes the AR-MA coefficients of model in the normal condition (before changing the bearing properties), and the red circle symbol shows the AR-MA coefficients after changing the boundary conditions. The two kinds of coefficients were overlaid in each figure for comparing the different distributions. In all figures, the first row plots the coefficient distributions of the tower points (#Pt1 and #Pt4) according to *x*-direction, and the second row shows the coefficient distributions of the girder points (#Pt2 and #Pt3) according to *z*-direction as well as the coefficient plot of the quarter span points (#Pt3) along to *x*-direction.

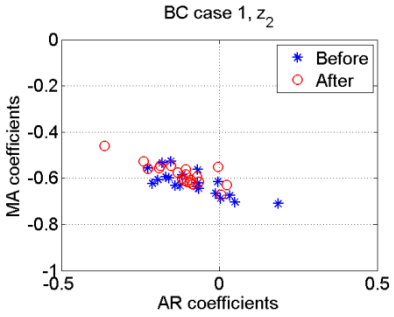


(a) #Pt1 – x direction

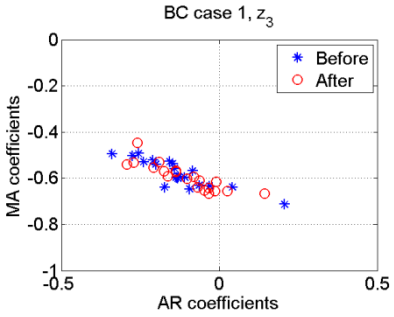


(b) #Pt4 – x direction

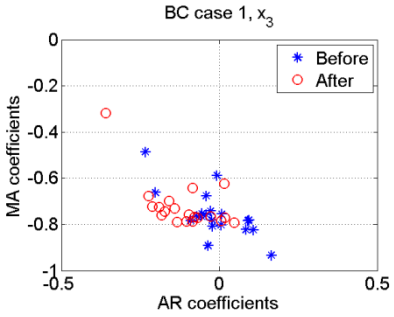




(c) #Pt2 – z direction

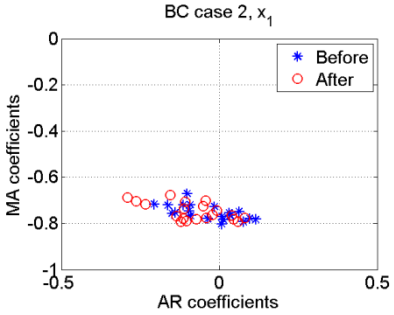


(d) #Pt3 – z direction

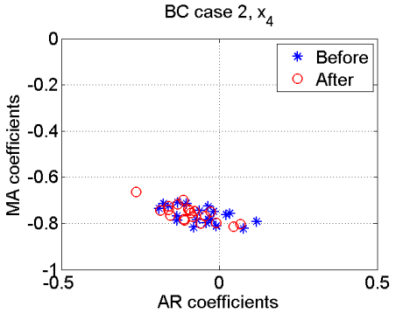


(e) #Pt3– x direction

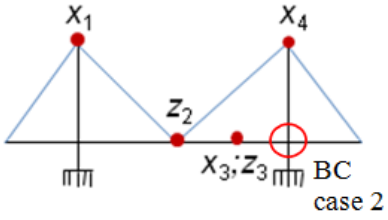
Figure 4.23 AR-MA coefficient plots of the BC case 1

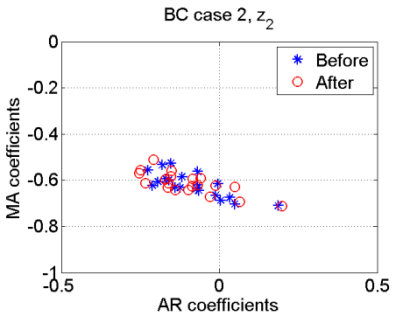


(a) #Pt1 – x direction

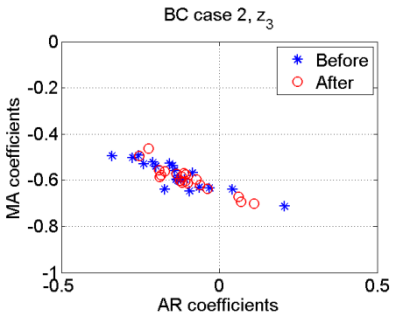


(b) #Pt4 – x direction

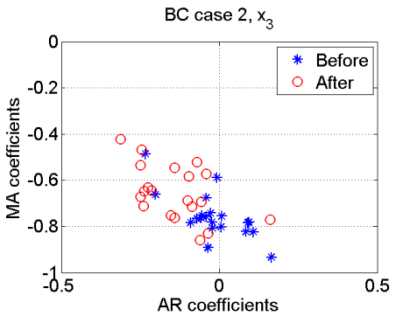




(c) #Pt2 – z direction

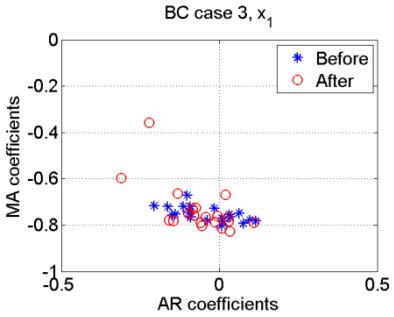


(d) #Pt3 – z direction

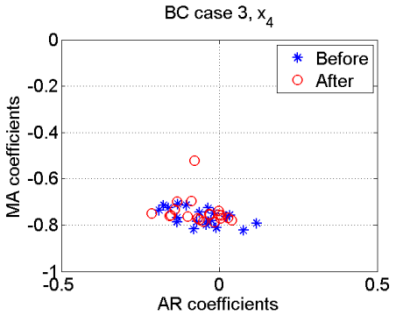


(e) #Pt3– x direction

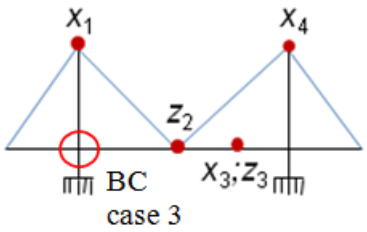
Figure 4.24 AR-MA coefficient plots of the BC case 2

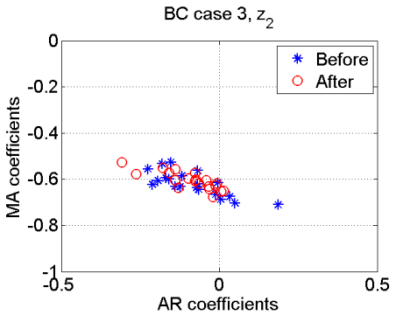


(a) #Pt1 – x direction

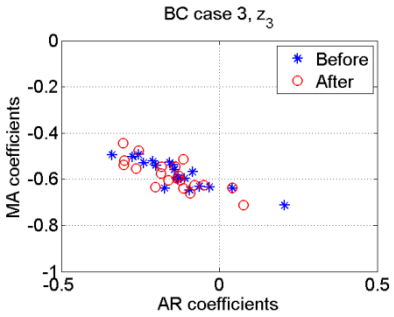


(b) #Pt4 – x direction

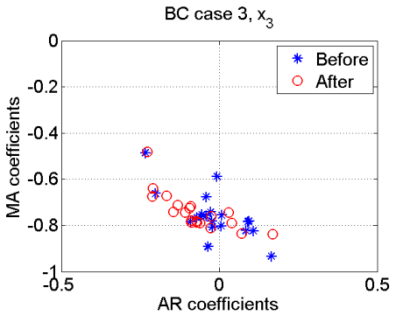




(c) #Pt2 – z direction

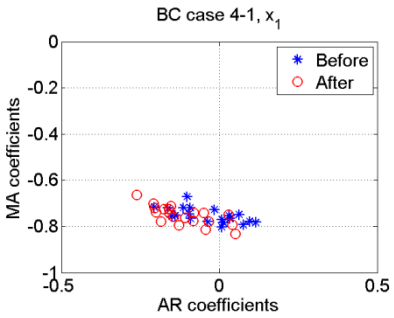


(d) #Pt3 – z direction

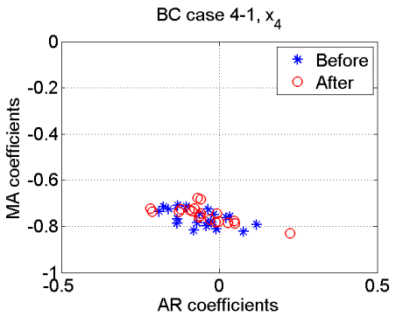


(e) #Pt3– x direction

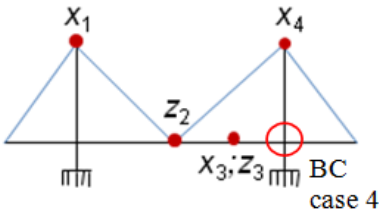
Figure 4.25 AR-MA coefficient plots of the BC case 3

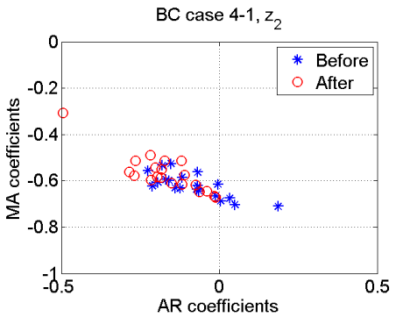


(a) #Pt1 – x direction

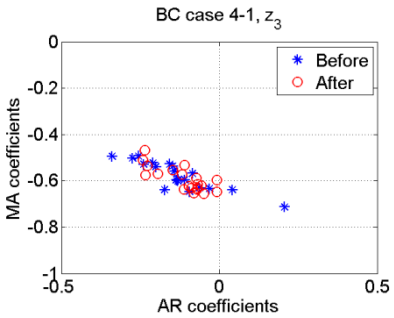


(b) #Pt4 – x direction

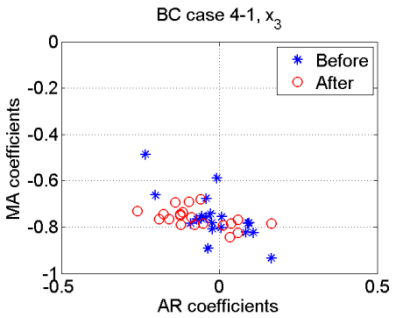




(c) #Pt2 – z direction

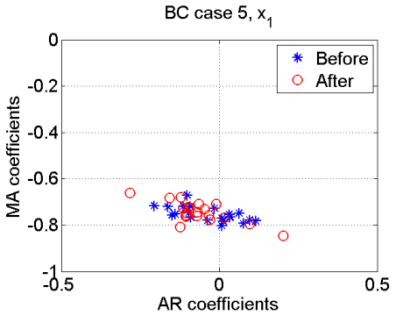


(d) #Pt3 – z direction

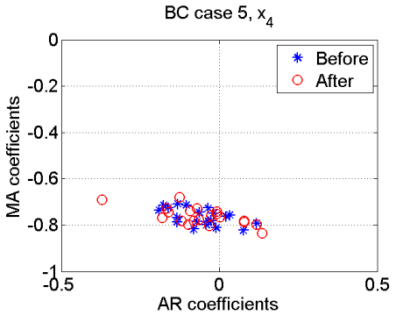


(e) #Pt3– x direction

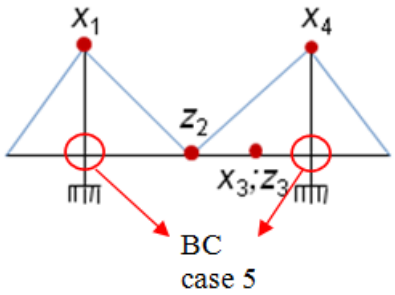
Figure 4.26 AR-MA coefficient plots of the BC case 4

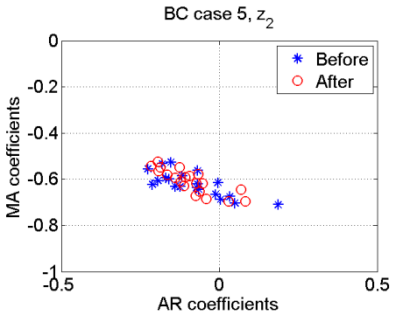


(a) #Pt1 – x direction

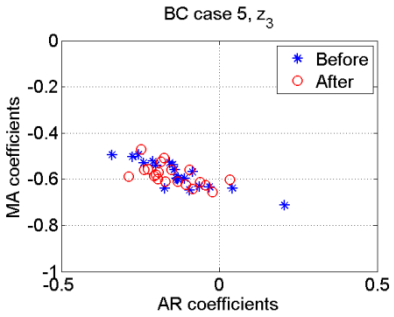


(b) #Pt4 – x direction

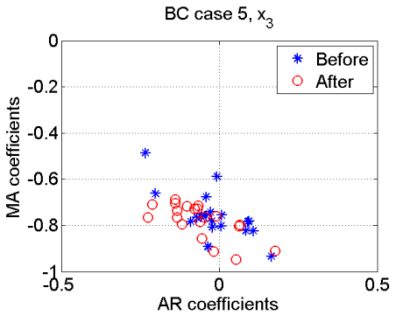




(c) #Pt2 – z direction

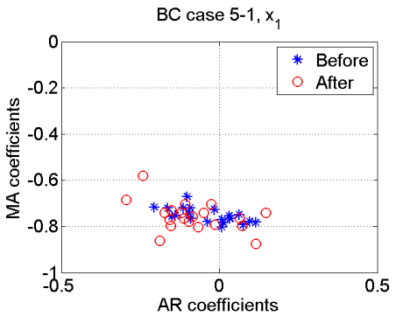


(d) #Pt3 – z direction

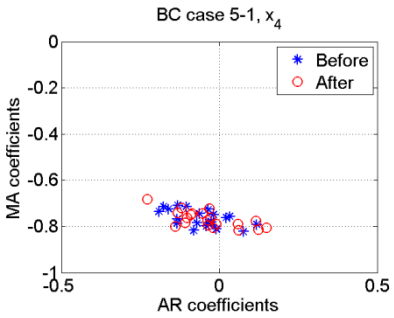


(e) #Pt3– x direction

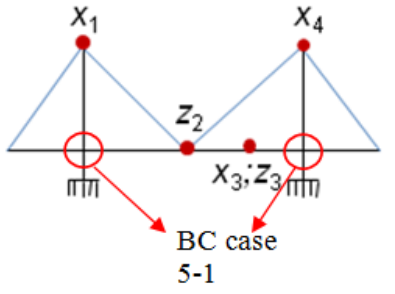
Figure 4.27 AR-MA coefficient plots of the BC case 5

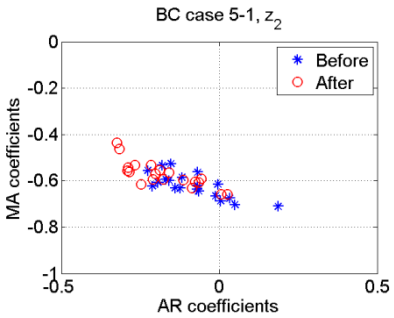


(a) #Pt1 – x direction

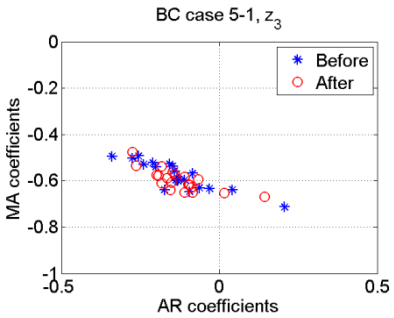


(b) #Pt4 – x direction

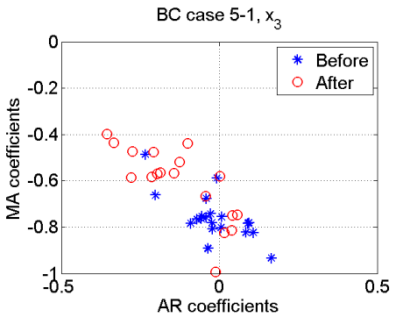




(c) #Pt2 – z direction



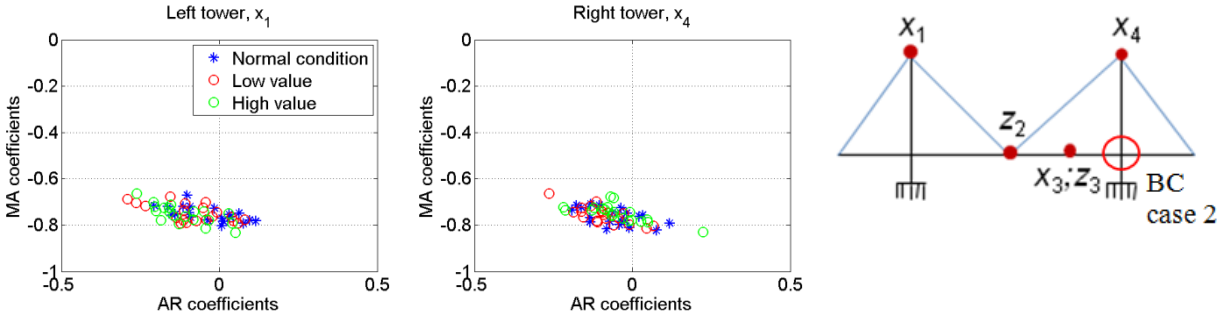
(d) #Pt3 – z direction



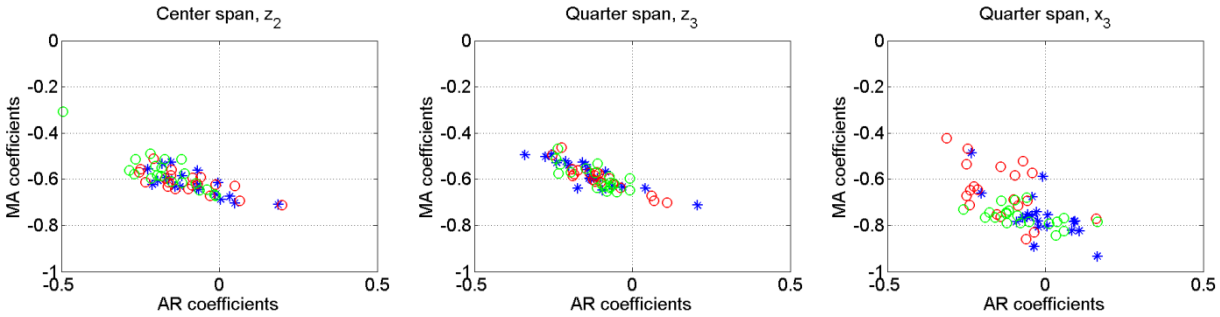
(e) #Pt3– x direction

Figure 4.28 AR-MA coefficient plots of the BC case 5-1

Furthermore, the AR-MA coefficients were extracted from the cases that have the same location but changed values from low to high of boundary conditions (BC case 2 and BC case 4; BC case 5 and BC case 5-1) are overlaid in the same figures for comparison, as shown in Figure 4.29 and Figure 4.30.

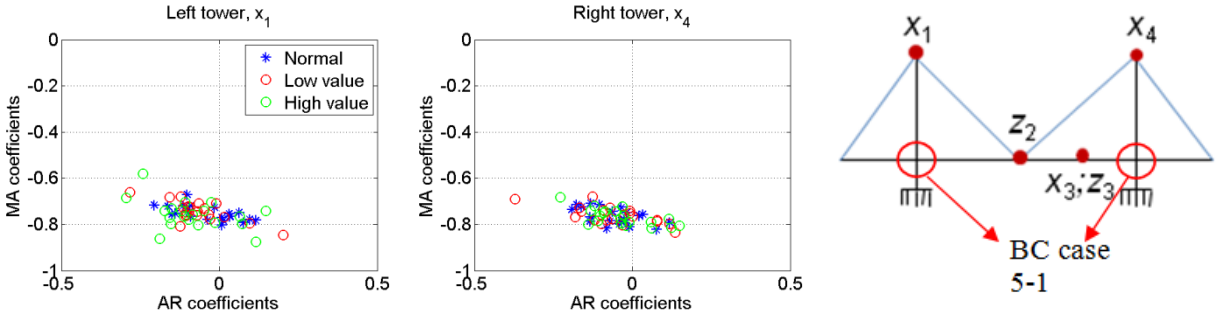


(a) #Pt1 – x direction (b) #Pt4 – x direction

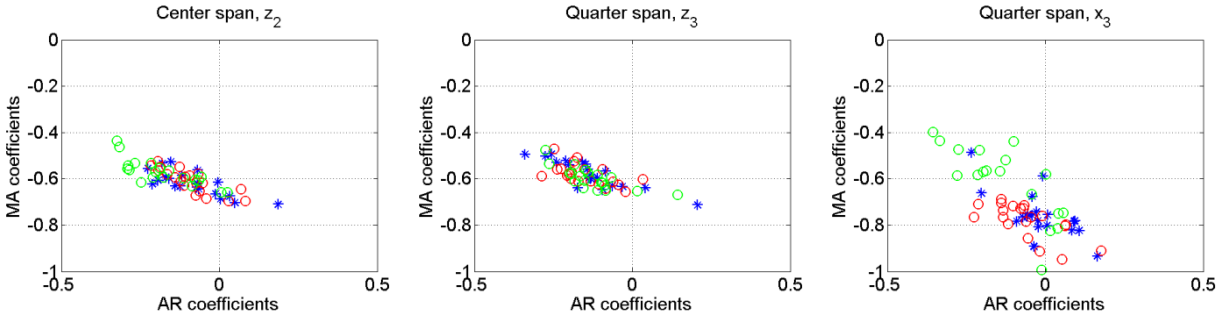


(c) #Pt2 – z direction (d) #Pt3 – z direction (e) #Pt3– x direction

Figure 4.29 Overlaid plot of AR-MA coefficients in the cases of BC case 2 and BC case 4



(a) #Pt1 – x direction (b) #Pt4 – x direction



(c) #Pt2 – z direction (d) #Pt3 – z direction (e) #Pt3– x direction

Figure 4.30 Overlaid plot of AR-MA coefficients in the cases of BC case 5 and BC case 5-1

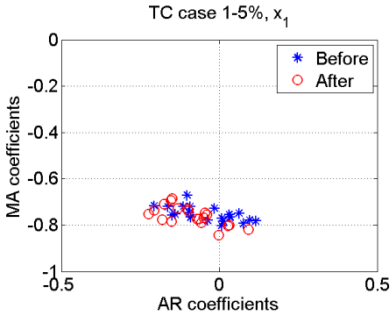
4.4.1.2Discussion

Firstly, when the boundary conditions change, there is a different trend in the distribution of AR-MA coefficients to compare with the normal condition. However, this kind of trend is not so clear in almost significant directions except the *x*-direction of the quarter span point, even though there are some scatter coefficients. In the all plots of the quarter span point along to longitudinal direction, there is clearly difference between the distributions of coefficients in two cases. It can be explained that, its location is the non-symmetric, thus the boundary condition changes can be easier to realize there by looking at the AR-MA coefficient distributions. In addition, the Figure 4.29 is the coefficient plots of two cases of bearing property changes in the same location (the right tower). Looking at those plots, the coefficient distributions seem to change when the bearing properties increase. This kind of trend in distribution could be easily realized in the *x*-direction of #Pt1, #Pt3 and the *z*-direction of #Pt2 (center span point). This point of discussion looks the same in the cases of the Figure 4.30 (BC case 5 and BC case 5-1). Overall, even though there is not so clear difference in the distributions of the AR-MA coefficients when the boundary condition changes, the trend in their distributions could be used to understand the structural condition change.

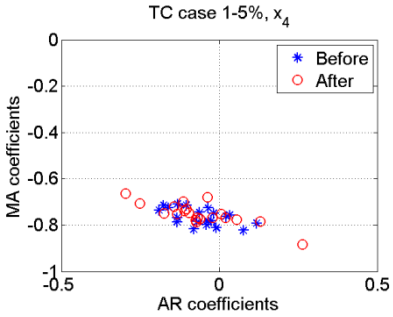
4.4.2. The case of cable tension changes

4.4.2.1 Plots of AR-MA coefficients

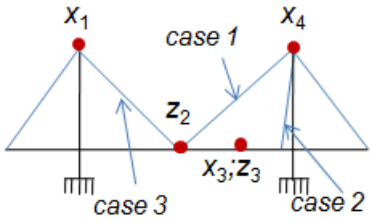
Similar to the case of boundary condition change, the coefficients of the normal condition and the cable tension reductions were figured out from Figure 4.31 to Figure 4.38, continuously. Especially, the tension change in case 1 were investigated in four cases of tension reductions which are 5%, 10%, 25% and 50% as shown from the Figure 4.31 to Figure 4.34, and Figure 4.35 shows the overlaid AR-MA coefficients in three cases of tension reduction (-5%, -25%, -50%) for comparison. The tension change in case 2 and case 3 are shown in the Figure 4.36 and Figure 4.37, respectively. Figure 4.38 shows comparison in AR-MA coefficient distributions between three selected cables in case of tension reduction of 10%.

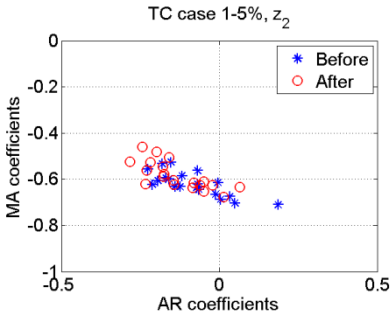


(a) #Pt1 – x direction

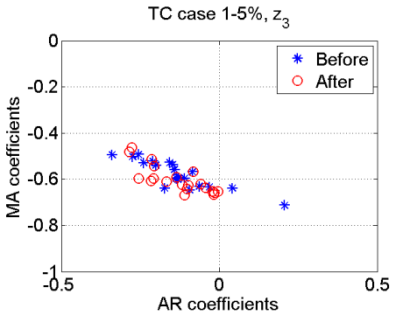


(b) #Pt4 – x direction

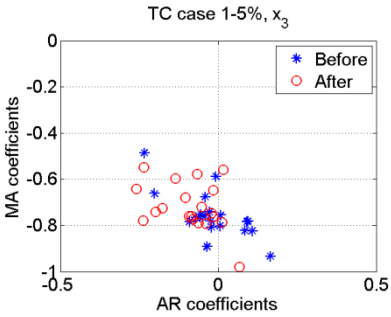




(c) #Pt2 – z direction

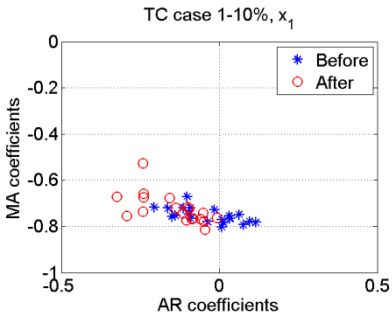


(d) #Pt3 – z direction

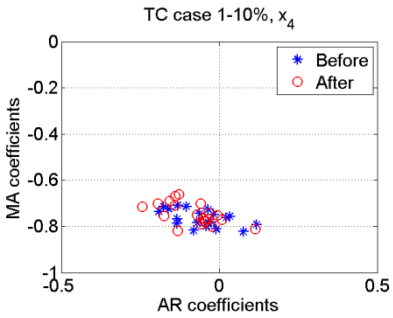


(e) #Pt3– x direction

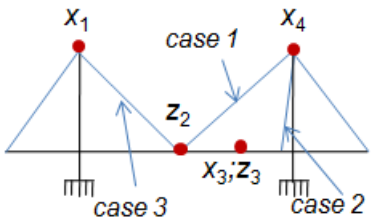
Figure 4.31 AR-MA coefficient plots of the tension change in case 1 (5% reduction)

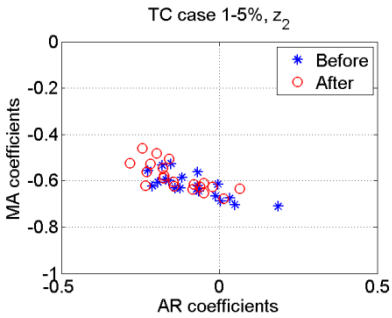


(a) #Pt1 – x direction

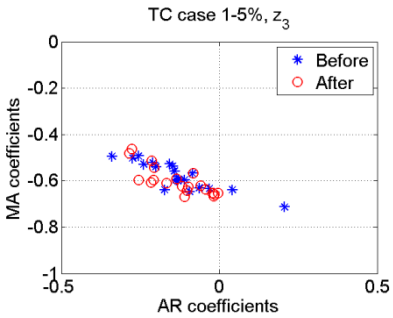


(b) #Pt4 – x direction

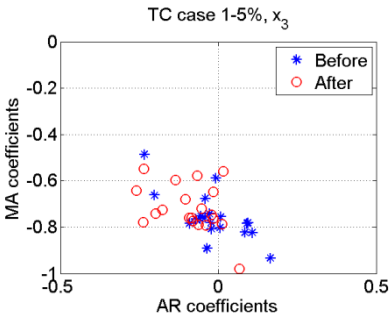




(c) #Pt2 – z direction

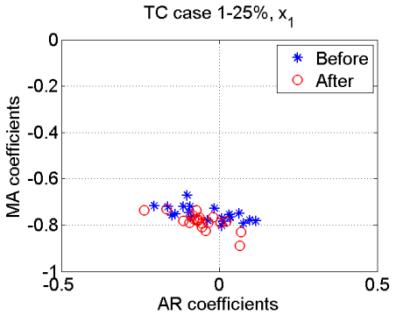


(d) #Pt3 – z direction

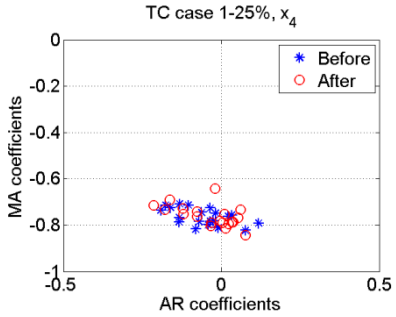


(e) #Pt3– x direction

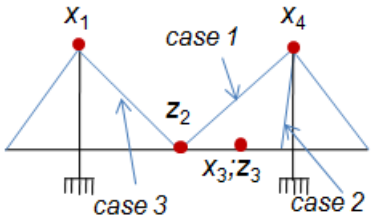
Figure 4.32 AR-MA coefficient plots of the tension change in case 1 (10% reduction)

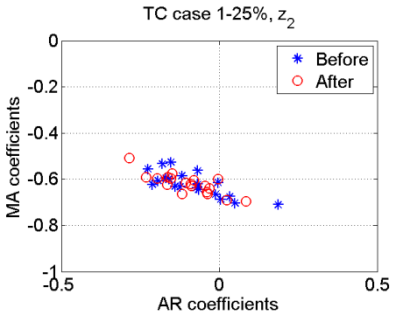


(a) #Pt1 – x direction

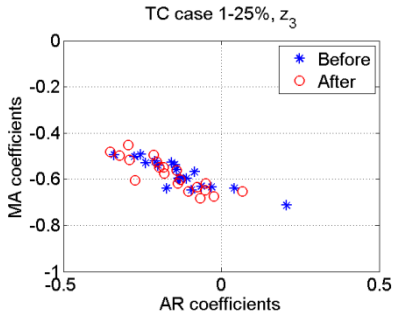


(b) #Pt4 – x direction

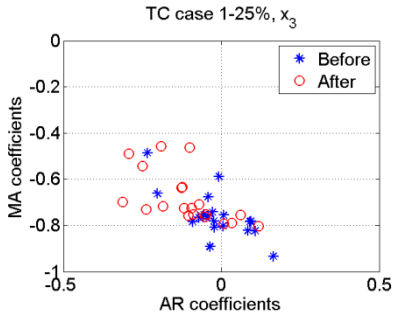




(c) #Pt2 – z direction

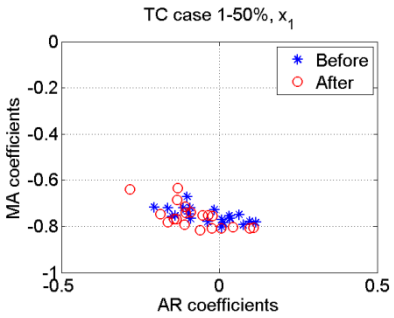


(d) #Pt3 – z direction

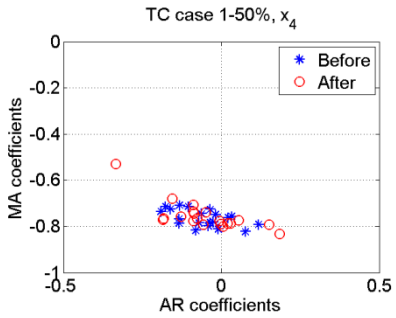


(e) #Pt3– x direction

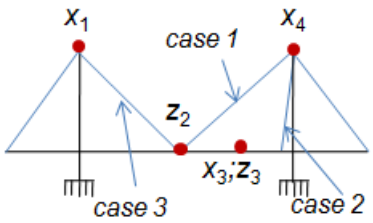
Figure 4.33 AR-MA coefficient plots of the tension change in case 1 (25% reduction)

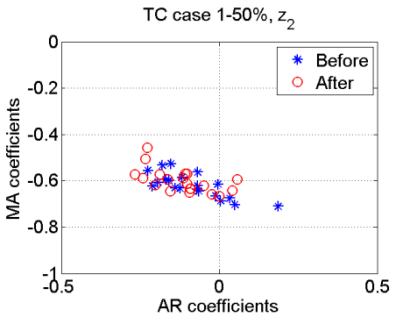


(a) #Pt1 – x direction

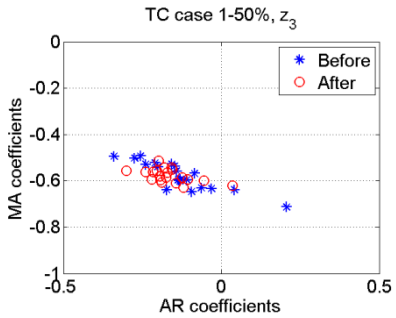


(b) #Pt4 – x direction

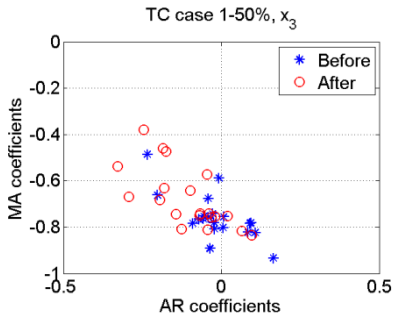




(c) #Pt2 – z direction

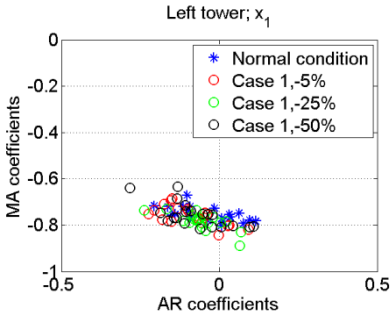


(d) #Pt3 – z direction

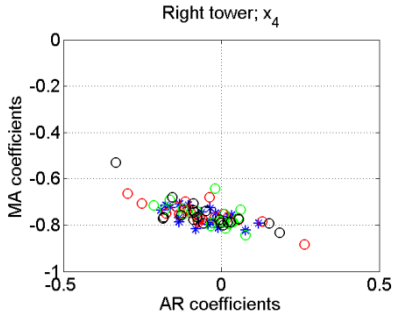


(e) #Pt3– x direction

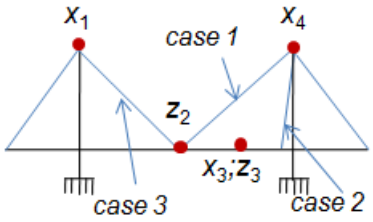
Figure 4.34 AR-MA coefficient plots of the tension change in case 1 (50% reduction)

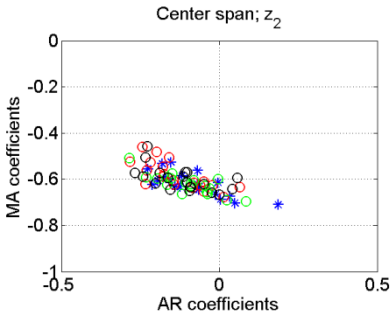


(a) #Pt1 – x direction

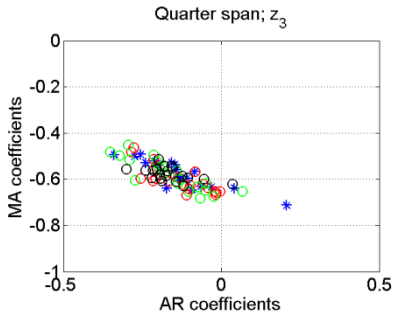


(b) #Pt4 – x direction

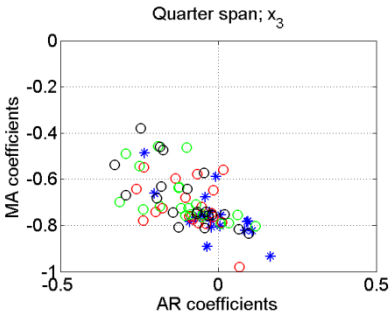




(c) #Pt2 – z direction

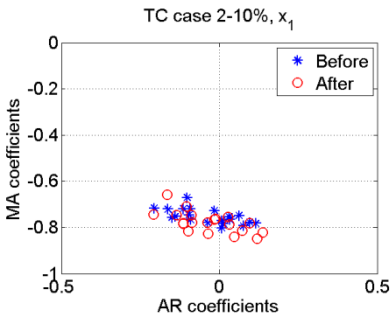


(d) #Pt3 – z direction

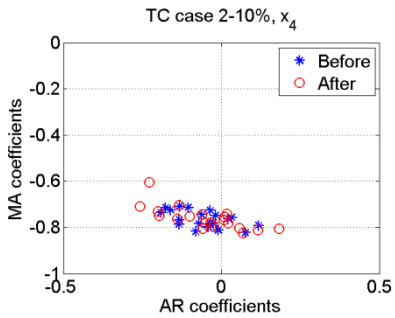


(e) #Pt3– x direction

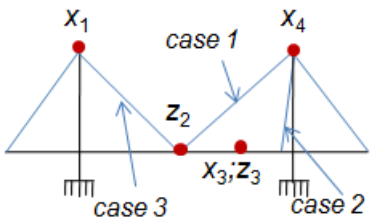
Figure 4.35 Overlaid AR-MA coefficient plots of the tension change in case 1

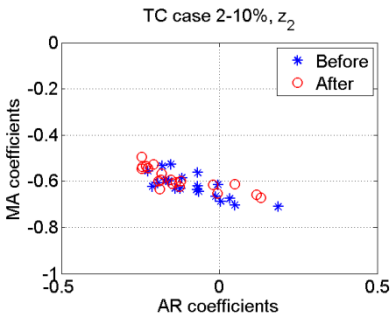


(a) #Pt1 – x direction

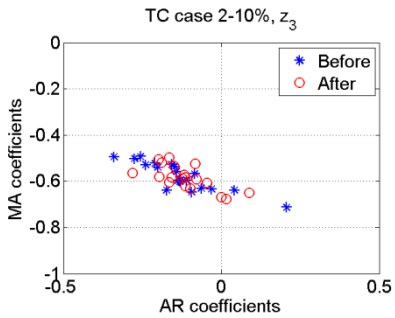


(b) #Pt4 – x direction

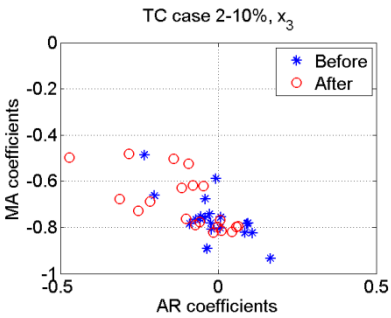




(c) #Pt2 – z direction

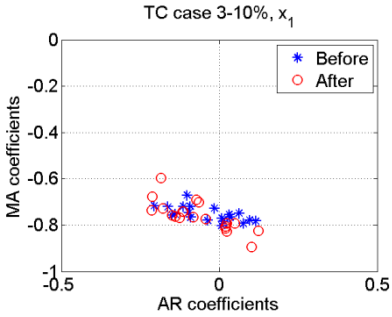


(d) #Pt3 – z direction

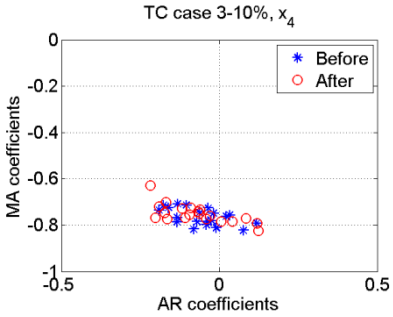


(e) #Pt3– x direction

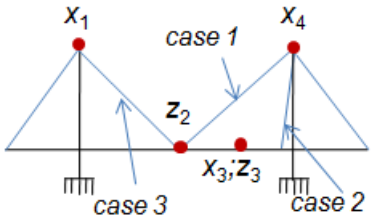
Figure 4.36 AR-MA coefficient plots of the tension change in case 2 (10% reduction)

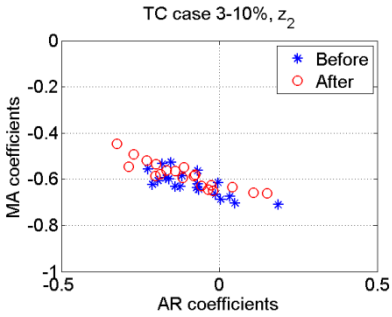


(a) #Pt1 – x direction

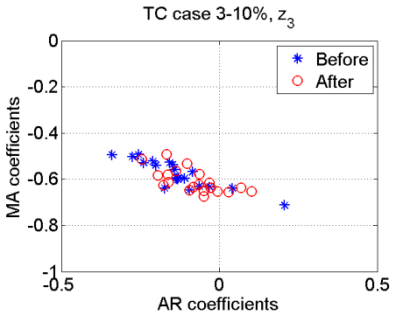


(b) #Pt4 – x direction

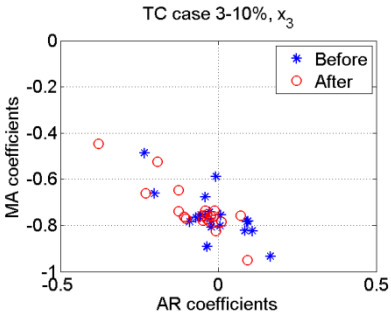




(c) #Pt2 – z direction

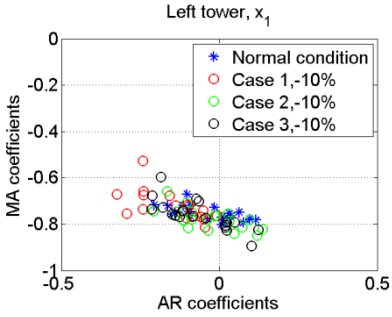


(d) #Pt3 – z direction

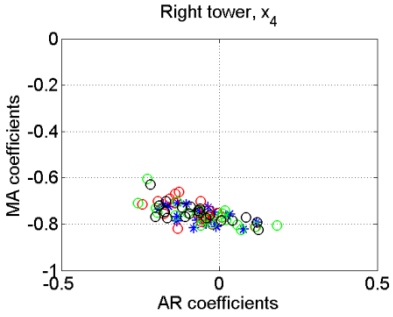


(e) #Pt3– x direction

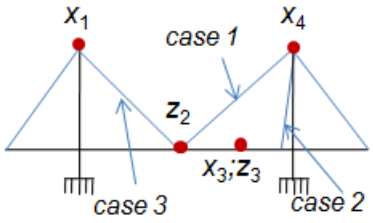
Figure 4.37 AR-MA coefficient plots of the tension change in case 3 (10% reduction)

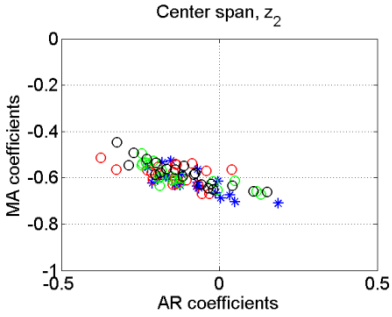


(a) #Pt1 – x direction

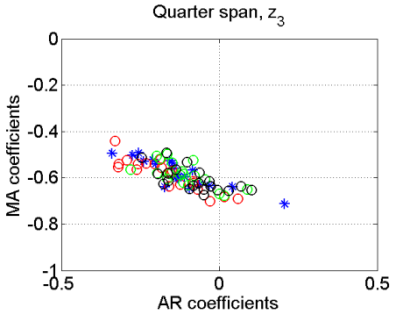


(b) #Pt4 – x direction

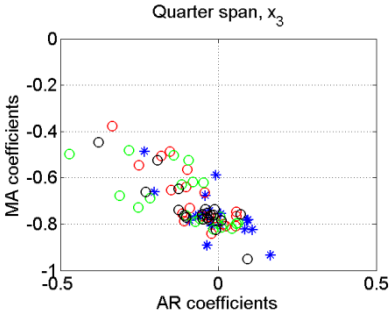




(c) #Pt2 – z direction



(d) #Pt3 – z direction



(e) #Pt3– x direction

Figure 4.38 Overlaid AR-MA coefficient plots of the tension change in three cases (-10%)

4.4.2.2Discussion

The distribution of AR-MA coefficients in the normal condition are used as the base for comparison with its distributions when the tension of cable decreases. The point of discussion here looks the similar to the case of boundary condition changes, that there is not so clear in the distribution of the coefficients in the almost significant directions except the *x*-direction of the quarter span point. The cable in case 1 was choose to investigated in four cases of tension reduction, because this cable has the highest tension, and the change in tension may clearly affect to the structural conditions. However, the distribution of AR-MA coefficients in these cases did not change so much to compare with the normal case. Another comparison in the cases of tension reduction of 10% (see Figure 4.38), even though the trend in coefficient distribution changes in each case, this trend is not so clear except the case of *x*-direction of the quarter span point. Overall, the similar concluding remark here is that the distribution of the AR-MA coefficients could be used to realize the structural condition changes when the tension of cable changes. However, there is not so different in distribution of coefficients in two cases of the normal condition and the structural condition changes. Therefore, the Mahalanobis distance method needs to use for the AR-MA coefficients to detect the structural condition changes. The application of Mahalanobis distance method is mentioned in the next chapter.

4.5. Summary

In this chapter, a cable-stayed bridge model is built in the MIDAS civil software to validate its global deformation behaviors under the effects of one-day time-series temperature changes as well as the ability of using AR-MA coefficients to detect the structural condition changes. The section 4.2 shows the description of the numerical cable-stayed bridge model as well as the input time-series temperature data. The explanation of the case studies of the structural condition changes in detail are also shown in this section. The acquired time-series displacements of the interested points with adding random noise are displayed in the end of this section. Then, the application of the ARIMA(1,1,1) model for time-series displacement data is shown (section 4.3), and the AR-MA coefficients are figured out to show their trend in distribution in section 4.4. The important conclusion in this chapter is that, the validation of trend in distribution of the AR-MA coefficients could be used to realize the structural condition changes, and those coefficients could be also used as multivariate features for outlier detection algorithm in the next chapter.

Chapter 5

Multivariate feature comparison using Mahalanobis distance method

5.1. Introduction

In chapter 3, an ARIMA model is applied for all GPS time series data of one year 2013 to extract the AR-MA coefficients for showing the distribution plots of them. Despite the effects of missing data were ignored in this case of study, there are several coefficients that scattered in almost directions. They might were caused by noise and abnormal data because there is no significant changes of the target bridge that were observed. As mentioned in the previous chapter, any outlier detection procedure should be adopted for the statistical structural condition diagnosis, where the distribution of AR-MA coefficients would be used as a base for any outlier detection procedure to assess structural conditions.

Moreover, a numerical cable-stayed bridge model was built in the chapter 4 to validate the ability of using the AR-MA coefficients for the structural condition assessments when some structural conditions change. The distribution plots of the AR-MA coefficients in numerical analysis did not clearly show the differences when the boundary conditions or the cable tensions of the model changed. Therefore, the algorithm of outlier detection should be adopted to use apply for the multivariate AR-MA coefficients.

The idea here is to apply the Mahalanobis distance (MD) method for comparing the estimated AR-MA coefficients which were extracted from the GPS time-series data and the numerical time-series data. There are two kinds of coefficients that are divided in this case of study, base coefficient and estimated coefficients, respectively. For the case of AR-MA coefficients acquired from GPS time-series data, the base coefficients taken to compare that are one month, three months or six months of coefficients, respectively. For the case of numerical analysis, the base coefficient is the AR-MA coefficient that acquired from the case of normal condition, and the coefficients extracted from all the cases of the structural condition changes (boundary condition changes and cable tension changes) are estimated coefficients. And all

coefficients were then used to calculate the Mahalanobis distance to compare the global deformation patterns.

5.2. General of Mahalanobis distance method

Mahalanobis distance is a distance measure that was introduced by Mahalanobis P. C in 1936. Mahalanobis distance method is based on correlation between variables by which different patterns can be identified and analyzed. Considering an example with a sample P and a population D, Mahalanobis distance is a multi-dimensional generalization of the idea of measuring how many standard deviations away sample variable V is from the mean of population P. This distance is zero if V is at the mean of P, and grows as P moves away from the mean. It is an useful way of determining similarity of an unknown sample set to a known sample set. Mahalanobis distance is a background method that was used in various studies, and it is also useful method for outlier detection. For instance, Mahalanobis distance method is the background ideal to propose a new display to emphasize unmasking multivariate outliers and leverage points [43]. A procedure was also proposed for the detection of outliers in multivariate data based on the idea of Mahalanobis distance [44]. In case of developing to propose the robust distance method for outlier detection, the mathematical background of Mahalanobis distance was used as a key idea [42,45,46].

The Mahalanobis distance method had proposed for comparing the multivariate feature vectors and the algorithm of statistical structural damage detection in some studies [40-42,47]. That is the normalized distance from the mean, as general defined in:

$$MD_i = \sqrt{(\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})} \tag{5.1}$$

where: $\mathbf{x}_i$ is a multi-dimensional feature vector; $\bar{\mathbf{x}}$ and $\mathbf{S}$ are a mean vector and a covariance matrix. Therefore, the algorithm of Mahalanobis distance is applied for the base coefficient features can be similarly calculated by:

$$MD_j^B = \sqrt{(\mathbf{x}_j^B - \bar{\mathbf{x}}^B)^T \mathbf{S}_B^{-1} (\mathbf{x}_j^B - \bar{\mathbf{x}}^B)} \tag{5.2}$$

where: The subscript B denotes the base data sets

$\bar{\mathbf{x}}^B$ is the mean vector of obtained feature vectors x_j of the baseline

$\mathbf{S}_B$ is the estimated covariance matrix of the baseline

In the practical case, the base data sets were selected to be three cases; one month (January), three months (from January to March) and six months (from January to June) in 2013,

respectively. For the numerical case, the base data is the coefficients of the normal condition of model. The estimated coefficients in two cases are then calculated the Mahalanobis distance that use the mean values and the covariance matrices of baseline. Therefore, the Mahalanobis distances of the estimated coefficients are calculated by Equation 5.3 as follows:

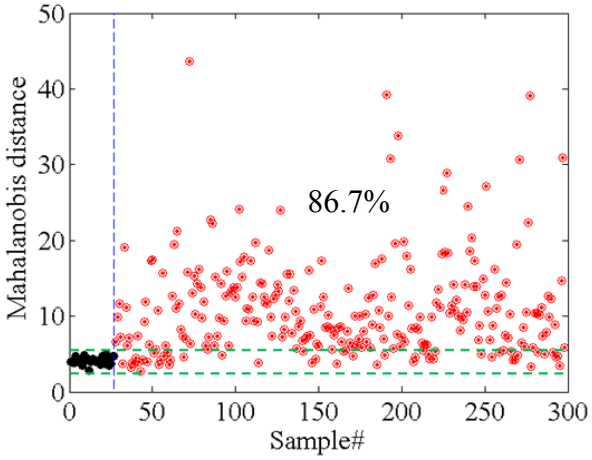
$$MD_k^E = \sqrt{(\mathbf{x}_k^E - \bar{\mathbf{x}}^B)^T \mathbf{S}_B^{-1} (\mathbf{x}_k^E - \bar{\mathbf{x}}^B)} \tag{5.3}$$

5.3. Application of Mahalanobis distance method

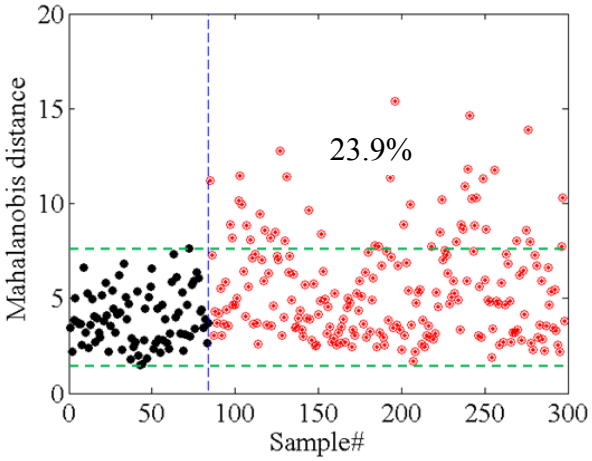
5.3.1. The case of AR-MA coefficients extracted from GPS time-series data (practical case)

The algorithm of Mahalanobis distance method was then applied for the AR-MA coefficients that extracted from almost one year GPS time-series data of the target bridge. For the base coefficient sets, all coefficients are arranged as a matrix coefficient that the column are sequentially arranged by AR and MA coefficients, and the rows are the sequence number of days. Figure 5.1 (a)-(c) are the plots of the Mahalanobis distance results (MDs) for the comparison between the base coefficient sets and the estimated coefficients. Notice that the black dots are the baseline results (MD^B) and the red dots are the estimated MDs (MD^E) of the estimated coefficients. By seeing those figures, the MD^B results of one month baseline are smaller to compare with the two other cases; i.e., the maximum value of one-month base is 4.92 to compare with 7.64 and 8.34 of the three-month base and six-month base, respectively. However, the MD^E results in the case of one-month base are much larger than those results in the two other cases. Additionally, there is a large difference between the MD^B and the MD^E of the one-month case. When the size of baseline increases, such as the three-month case and the six-month case, the differences between the MD^B and the MD^E decrease. It means that, the size of baseline influenced on the distributions of the MD^E.

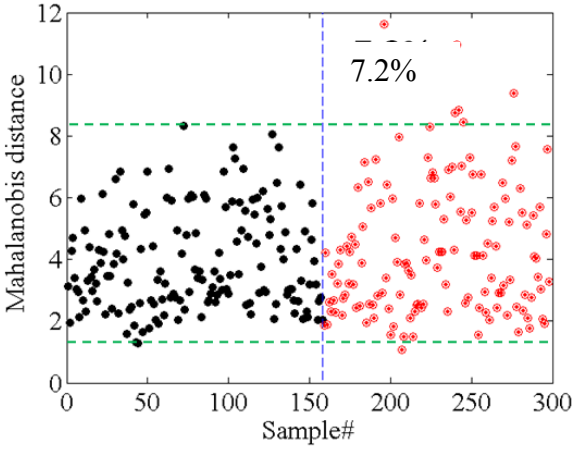
The effect of the size of baseline was investigated by taking the base ranges that are limited by the maximum and the minimum values of the MD^B. The results show that there is 86.7% of MD^E of the one-month case which is the outlier, while these results of the three-month and six-month cases are 23.9% and 7.2%, respectively. In this target bridge, it was confirmed that there was no significant changes in the structural condition by the inspection. Therefore, the baseline distribution from the six-month data is appropriate to construct the outlier detection algorithm of the structural condition assessment. Moreover, it was also considered that the baseline distribution should be carefully defined in relation to the actual structural condition, and it would be better to take as large number of base data sets as possible.



(a) Based on one-month baseline



(b) Based on three-month baseline



(c) Based on six-month baseline

Figure 5.1 Mahalanobis distance plots (black dots: MD^B baseline; red circles: estimated MD^E distance)

5.3.2. The case of AR-MA coefficients extracted from numerical analysis

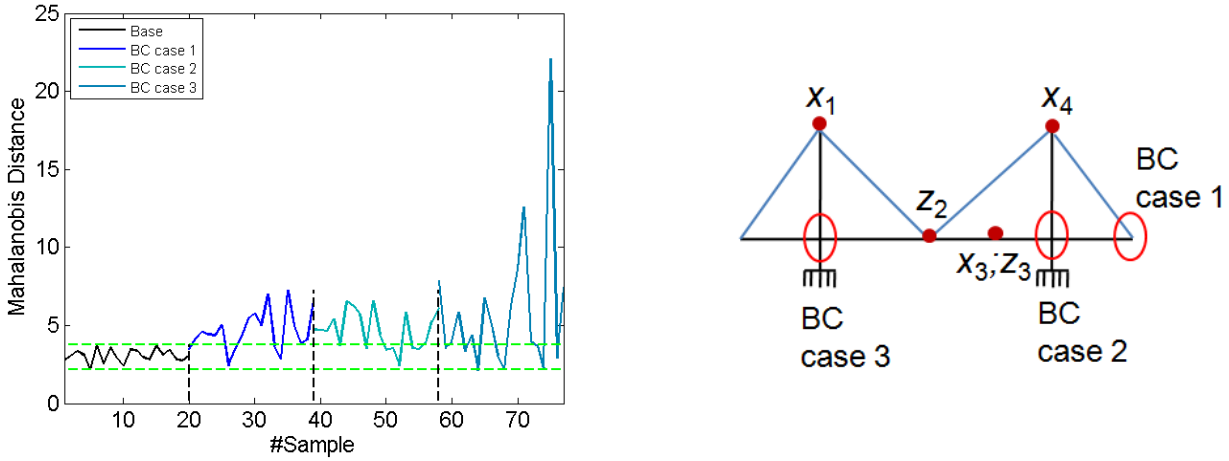
The Mahalanobis distance method was applied for all AR-MA coefficients that extracted from the numerical time-series data. Here, the base coefficient is the AR-MA coefficients of the normal condition. The Mahalanobis distance of base coefficients are figured out by the black line in all figures, and the base range is generated in the green lines by taking the maximum and minimum values of the MD^B. The mean values and the covariance matrix estimated from the base coefficients were used to calculate the Mahalanobis distance of the estimated coefficients. Then, the estimated Mahalanobis distance results are plotted in the different colors of different cases.

5.3.2.1 Comparison of Mahalanobis distance results in the case of boundary condition changes

Firstly, the Figure 5.4 shows the comparison of the MDs results between the base and the estimated MDs, as well as the comparison between three cases of the boundary condition changes. It can be seen that, there is a difference in the MD^E values of three cases: BC case 1, BC case 2 and BC case 3, that the MD^E values of the case 3 are much larger than case 1 and case 2. This point of view sounds a little strange, because the location of case 3 is farther than case 1 and case 2 where are near the three interested point. Logically, the effect of case 3 should be less than case 1 and case 2. However, looking at the distribution of the AR-MA coefficients in the previous chapter (Figure 4.23, Figure 4.24 and Figure 4.25), it can be clearly seen that there are some outlier coefficients in the plots of tower points according to *x*-direction in BC case 3 (Figure 4.25). Therefore, the effect of those outlier coefficients is clearly displayed in the Mahalanobis distance results. In addition, there is a large amount of MD^E values in all three cases that are out of the base range. Statistically, there is 75% of the MD^E of case 1 that are out of base range, while those portions of case 2 and case 3 are 65% and 70 %, respectively.

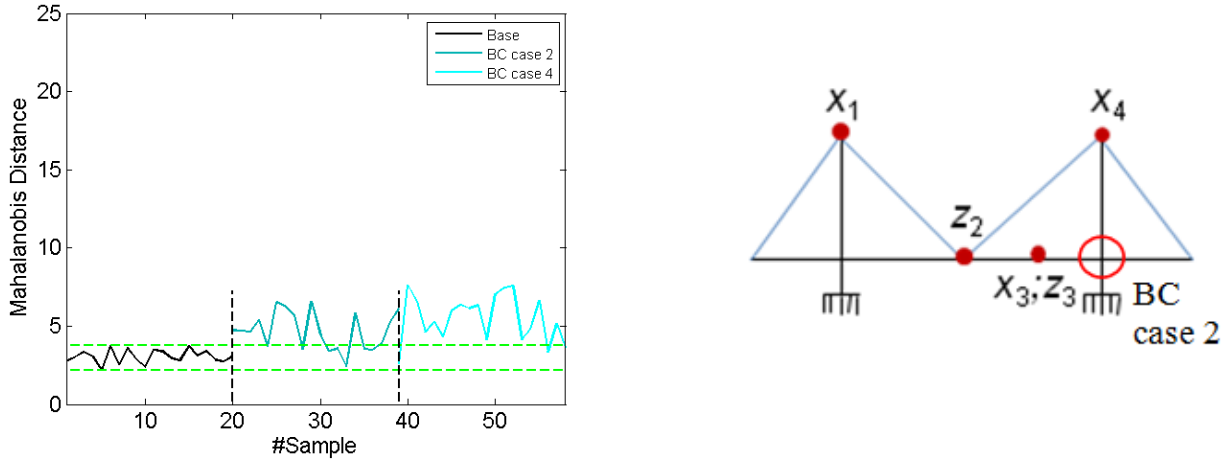
Secondly, the Figure 5.5 shows the comparison of the Mahalanobis distance results between two cases of boundary condition changes in the same location (BC case 2 and BC case 4). The bearing properties of case 4 are much higher than the case 2. Looking at the Figure 5.5 (a), the MD^E values of the case 4 are also larger than the case 2. Statistically, there are 75% of the MD^E results of case 4 that are out of the base range, and this portion is also larger than the case 2. It can be understood that, when the bearing properties of structure increase to be very high, the structural condition changes can be also realized in the results of Mahalanobis distance. This kind of conclusion is also satisfied in the cases of results in the Figure 5.6, where compare the Mahalanobis distance results in two cases: BC case 5 and BC case 5-1 in the same location. In

detail, the bearing properties of case 5-1 are much higher than the case 5 and the Figure 5.6 (a) shows that the MD^E values of the case 5-1 are larger than case 5. Statistically, there are 80% of the MD^E results of case 5-1 that are out of the base range, and this portion is also larger than the case 5 (70%). It could be concluded that, the boundary condition changes of the model can be clearly realized in the plots of Mahalanobis distance results.



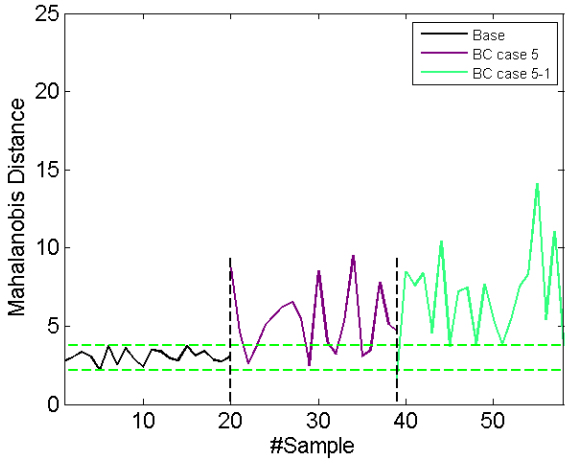
(a) Mahalanobis distance plots (b) Locations of boundary condition changes

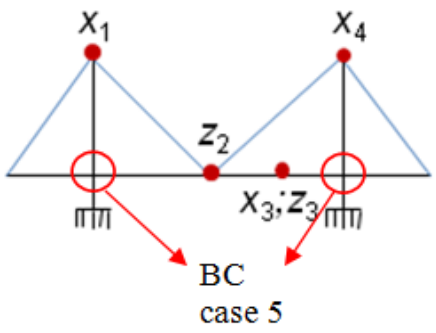
Figure 5.4 Comparison of Mahalanobis distance results in three different locations of boundary condition changes



(a) Mahalanobis distance plots (b) Locations of boundary condition changes

Figure 5.5 Comparison of Mahalanobis distance results between the boundary condition change in case 2 and case 4 (same location)





(a) Mahalanobis distance plots (b) Locations of boundary condition changes

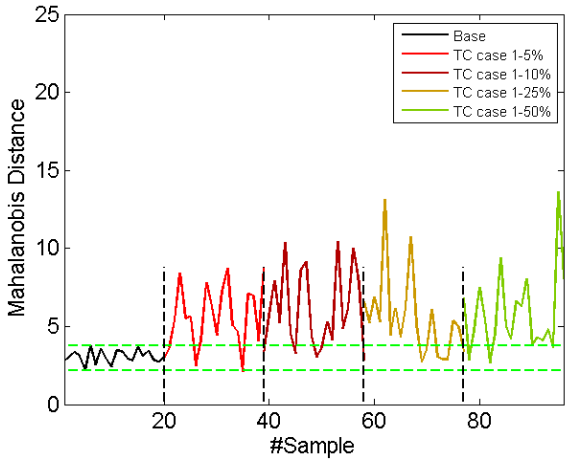
Figure 5.6 Comparison of Mahalanobis distance results between the boundary condition change case 5 and case 5-1 (same location)

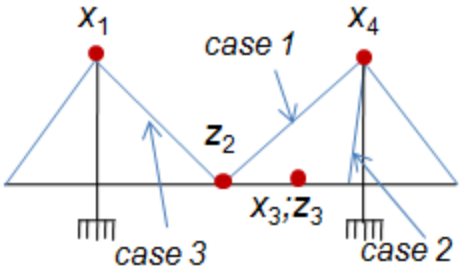
5.3.2.2Comparison of Mahalanobis distance results in the case of cable tension changes

First of all, the cable tension in case 1 was choose to change the tension in four reduction levels such as 5%, 10%, 25% and 50%. This is the one of the highest tensions, and its location is near almost the interested points. Therefore, the effects of tension reduction of case 1 need to be studied. The Figure 5.7 (a) shows the comparison of the Mahalanobis distance results of case 1 in four levels of the tension reductions. It can be seen that, the maximum MD^E value tends to increase when the tension of cable reduces. Additionally, there is a large amount of the MD^E values that are out of the base range. In statistically, there are 75% of the MD^E values that out of the base range in all three reduction levels (5%, 10% and 25%). This portion reaches to 85% when the tension reduces of 50%. It can be understood that, the application of Mahalanobis distance method for the AR-MA coefficients that can realize the reduction of cable tension.

Furthermore, the Figure 5.8 (a) shows the comparison of the Mahalanobis distance values in three cases of the cable tension changes at 10% (case 1, case 2 and case 3). The selected cable tension in case 3 is also one of the highest tension same as the case 1, and the selected cable tension in case 2 is one of the lowest tensions of the model. The purpose of this test is that, the tension reduction in different location of model may cause the different MD^E results. Looking at the Figure 5.8 (a), the MD^E values of the case 1 is larger than other cases, while the MD^E values of the case 2 is also larger than the case 3 even though the cable there has the lowest tension. It can be explained that the location of cables in case 1 and case 2 are near the three interested

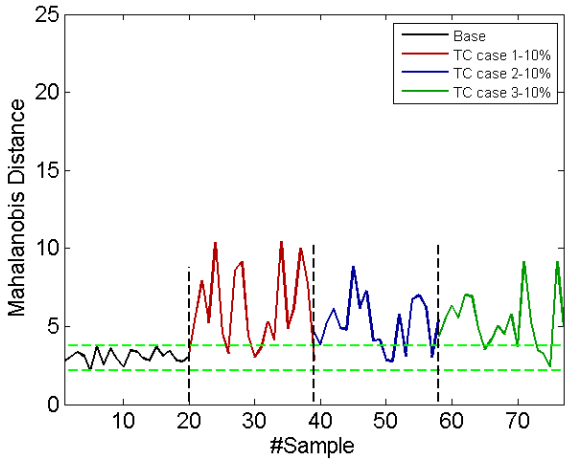
points, therefore their tension reductions have more effects than the case 3. Overall, similar to the case of boundary condition changes, the application of the Mahalanobis distance method for the AR-MA coefficients could realize the structural condition changes under the effects of temperature changes.

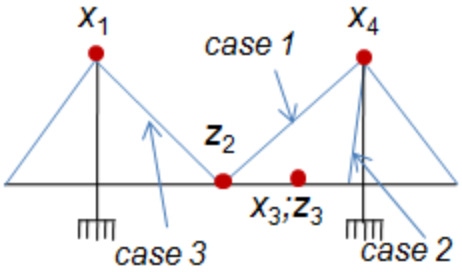




(a) Mahalanobis distance plots (b) Locations of boundary condition changes

Figure 5.7Comparison of Mahalanobis distance results of cable tension change in case 1 at four reduction levels (5%, 10%, 25%, and 50%)





(a) Mahalanobis distance plots (b) Locations of boundary condition changes

Figure 5.8Comparison of Mahalanobis distance results in three cases of cable tension change at 10% of tension reduction

5.4. Summary

This chapter studies on the application of the Mahalanobis distance method for comparing the multivariate AR-MA coefficients in both practical case and numerical case. The general Mahalanobis distance method is mentioned in the chapter 5.2. Then, the applications of Mahalanobis distance are shown in the chapter 5.3. The concluding remarks in this chapter is that; firstly, a half of year of the AR-MA coefficients should be taken as the base coefficient to compare with the new coefficient to realize the structural condition by using the Mahalanobis distance method. And the application of Mahalanobis distance method for the extracted AR-MA coefficients in the numerical analysis could be realized the structural condition changes.

Chapter 6

Concluding Remarks and Future Research

6.1. Concluding Remarks

The first part of this dissertation studies on the analysis of GPS time-series long-term data acquired from a practical long-span cable-stayed bridge. Then, the applicability of time-series analysis for the GPS long-term data considering the global deformation patterns is the aim for the structural condition assessment. Furthermore, a cable-stayed bridge model was built in the numerical analysis step to validate the ability of using AR-MA coefficients for the structural condition assessments. The results of this study show that:

- The correlation coefficients analysis was conducted using the pre-processed GPS data with the interpolation of missing data. The data in the significant directions at the two towers and the girder showed high correlation with the atmospheric temperature data.

- It was also shown that the GPS data with high correlation with the temperature dominated by one-day periodic behaviors. In addition, most of them were also highly correlated each other. The global deformation mode due to the one-day periodic temperature change could then be clarified, and the significant directions in displacements of the two towers and the girder were also shown.

- The applicability of estimating ARIMA model was verified to apply for the GPS long-term data. The results showed that the ARIMA(1,1,1) is the appropriate model for the GPS time-series data.

- The plots of extracted AR-MA coefficients showed the trend in their distributions, and they distributed in regression lines, especially in the significant directions that showed high correlation coefficients with the temperature data.

- In the numerical analysis, the distribution of the extracted AR-MA coefficients change when the boundary condition or the cable tensions change, even though there is a little different in their distribution in two cases; normal condition and structural condition changes.

- Both the AR-MA coefficients in two steps of study were then used as the multivariate features for comparison by applying the Mahalanobis distance method. The results showed that the baseline distribution for the outlier detection in practical case can be constructed by taking the estimated coefficients from the data sets acquired for a half of year to assess structural conditions under operation. The numerical results clearly show that the application of the Mahalanobis distance method for comparing the multivariate AR-MA coefficients could realize the structural condition change under operation such as the boundary condition change or the cable tension reduce.

Therefore, this study propose that, time-series analysis with the application of ARIMA model should be used to deal with a large amount of GPS long-term monitoring data to extract the AR-MA coefficients for understanding the structural conditions under operation. In the case of structural condition assessments using the AR-MA coefficients, the application of Mahalanobis distance method is necessary.

6.2. Recommendations for future works

In the structural health monitoring of bridges, there are many factors that influence on the health of structure, and those factors could come from the operational source or environmental source. This research only studied on the effects of the atmospheric temperature to the global deformation of whole structure through the analysis of GPS long-term monitoring data. In the future works, many other factors such as cross wind speed, loading, or local temperature change and so on that are recommended for study their influences on the health of structures through using time-series analysis. Furthermore, some algorithm of damage detection should carry out to validate for the practical study.

References

[1] *http://commons.wikimedia.org/wiki/File:Akashi_Kaikyo_Ohashi_01.jpg*

[2] *http://hongkong.panduanwisata.id/files/2011/10/tsing-ma-bridge2.jpg*

[3] *http://www.fhwa.dot.gov/publications/publicroads/11janfeb/03.cfm*

[4] Reza V. F (2013). Structural Health Monitoring and Damage Identification of Bridge Using Triaxial Geophones and Time Series Analysis. *Doctoral Dissertation*.

[5] Balageas D, Fritzen C. P, Guemes A, (2006). Structural Health Monitoring. *Book*, ISTE.

[6] Kaloop, M. R., & Li, H. (2009). Monitoring of bridge deformation using GPS technique. *KSCE Journal of Civil Engineering*, *13*(6), 423-431.

[7] Kaloop, M. R., & Li, H. (2011). Sensitivity and analysis GPS signals based bridge damage using GPS observations and wavelet transform. *Measurement*, *44*(5), 927-937.

[8] Celebi, M. (2000). GPS in dynamic monitoring of long-period structures. *Soil Dynamics and Earthquake Engineering*, *20*(5), 477-483.

[9] Sohn H, Dzwonczyk M, Straser E. G, Kiremidjian A. S, Law, K. H., & Meng T (1999). An experimental study of temperature effect on modal parameters of the Alamosa Canyon Bridge. *Earthquake engineering & structural dynamics*, *28*(8), 879-897.

[10] Cornwell P, Farrar C. R, Doebling S. W, & Sohn H (1999). Environmental variability of modal properties. *Experimental Techniques*, *23*(6), 45-48.

[11] Farrar C. R, Cornwell P, Doebling S. W, & Prime M. B, (2000). Structural Health Monitoring Studies of the Alamosa Canyon and I-40 Bridges. Los Alamos National Laboratory, LA-13635-MS.

[12] Cheng P, John W, & Zheng W (2002, April). Large structure health dynamic monitoring using GPS technology. In *FIG XXII International Congress, Washington, DC USA*.

[13] Fujino Y, Murata M, Okano S, & Takeguchi M, (2000). Monitoring system of the Akashi Kaikyo Bridge and displacement measurement using GPS. In *SPIE's 5th Annual International*

Symposium on Nondestructive Evaluation and Health Monitoring of Aging Infrastructure (pp. 229-236).International Society for Optics and Photonics.

[14] Ahmed El-Rabbany (2002). Introduction to GPS: The Global Positioning System. *Book*, Artech House.

[15] Sohn H, Czarnecki J. A, & Farrar C. R, (2000). Structural health monitoring using statistical process control. *Journal of Structural Engineering*, *126*(11), 1356-1363.

[16] Omenzetter P, & Brownjohn J. M. W, (2005). A seasonal ARIMAX time series model for strain-temperature relationship in an instrumented bridge. Proceedings of the 5th International Workshop on Structural Health Monitoring, 299-306.

[17] Omenzetter P, &Brownjohn J. M. W, (2006). Application of time series analysis for bridge monitoring. *Smart Materials and Structures*, *15*(1), 129-138.

[18] *http://news.bbc.co.uk/2/hi/asia-pacific/7013711.stm*

[19] *http://www.flickriver.com/photos/tags/canthobridge/interesting/*

[20] The design documentations of Can Tho cable stayed bridge (2010)

[21] Pigott T. D, (2001). A review of methods for missing data. *Educational research and evaluation*, *7*(4), 353-383.

[22] C. F. Gerald, P. O. Wheatley (2004). Applied Numerical Analysis. Seventh Edition, chapter 3, pp. 203-207.

[23] Ross Ihaka (2005). Time Series Analysis. Lecture Notes

[24] Shumway R. H, & Stoffer D. S, (2010). *Time series analysis and its applications: with R examples*. Springer.

[25] Hamilton J. D, (1994). Time Series Analysis. *Book*, New Jersey, USA: Princeton University Press.

[26] Sohn H, Czarnecki J. A, Farrar C. R, (2000). Structural Health Monitoring using statistical process control. *Journal of Structural Engineering*, 126(11): pp. 1356-1363

[27] Pigott T. D, (2001). A review of methods for missing data. *Educational Research and Evaluation*, 7(4): pp. 353-383.

[28] Hirose K, Kawano S, Konishi S, Ichikawa M, (2011). Bayesian information criterion and selection of the number of factors in factor analysis models. *Journal of Data Science*, 9(2): pp.243-259.

[29] Fuller W. A, (1976). Introduction to Statistical Time Series. New York: John Wiley and Sons. ISBN 0-471-28715-6.

[30] Elliott G, Rothenberg T. J, Stock J. H, (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrics* **64** (4): 813–836. JSTOR 2171846

[31] G. M. Ljung, G. E. P. Box, (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika* **65** (2): 297–303.doi:10.1093/biomet/65.2.297.

[32] Davidson, James (2000). Econometric Theory. Blackwell. p. 162. ISBN 0-631-21584-0.

[33] Tsay R. S, (2010). Analysis of Financial Time Series. 3rd ed. Hoboken, NJ: John Wiley & Sons, Inc., 2010.

[34] Box G. E. P. and D. Pierce (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*. Vol. 65, pp. 1509–1526.

[35] Konstantakopoulos T. G, Michaltsos G. T, (2010). A mathematical Model for a Combined Cable System of Bridges. *Engineering Structures* ; 32(9):2717-2728. DOI: 10.1016/j.engstruct.2010.04.042

[36]. Ozlem C, Alemdar B, Suleyman A (2010). Stochastic Finite Element Analysis of a Cable Stayed Bridge System with Varying Material Properties. Probabilistic Engineering Mechanics, 25: pp. 279-289.

[37]. Wu Q, Takahashi K, Nakamura S (2008). Response Characteristics of Local Vibration in Stay Cables on an Existing Cable Stayed Bridge. *International Conference on Bridge Maintenance, Safety and Management*, Seoul, Korea

[38]. Ronal M. M (2000). Corrosion in Suspension Bridge Cables. *16th Congress of IABSE*, Licerne.

[39]. Vikas A. C, Prashanth M. H, Indrani G, Channappa T. M (2013). Effect of Cable Degradation on Dynamic Behavior of Cable Stayed Bridges. Journal of Civil Engineering Research. 3(1): 35-45. DOI: 10.5923/j.jce.20130301.04

[40] Nishio, M. *et al* (2011). Feature Extraction for Structural Dynamic Model Validation. Conference Proceedings of the Society for Experimental Mechanics Series, pp. 153-163.

[41] K. Worden and G. Manson (2000). Damage Detection Using Outlier Analysis. *Journal of Sound and Vibration*, 229(3): pp.647-667.

[42] Peter F., Robert G. G, Clemenn R, (2005). Multivariate outlier detection in exploration geomistry. *Computers & Geosciences* 32(2005): pp.579-587.

[43] Peter J. R, Bert C. v. Z, (1990). Unmasking Multivariate Outliers and Leverage Points. *Journal of the American Statistical Association*, 85(411): pp. 633-639

[44] Ali S. H, (1992). Identifying Multiple Outliers in Multivariate Data. *Journal of the Royal Statistical Society B (Methodological)*, 54(3): pp. 761-771

[45] Peter J. R, Katrien V. D, (1999). A fast Algorithm for the Minimum Covariance Determinant *Estimator, Technometrics*, 41(3): pp. 212-223

[46] Donald A. J, Young C, (2004). Robust principal component analysis and outlier detection with ecological data. *Environmetrics*, 15: pp. 129-139

[47] Dionysius M. S, Tomonori N, Yozo F, Di S, Chondro T, Hirotaka M, (2013). Vibration study and Application of Outlier Analysis to the S101 Bridge Full-Scale Destructive Testing. *Industrial Safety and Life Cycle Engineering*. pp. 305-330

[48] Ruigen Y, Shamim N. P, (2012). Autoregressive statistical pattern recognition algorithms for damage detection in civil structures. *Mechanical Systems and Signal Processing*, DOI: 10.1016/j.ymssp.2012.02.014